STIC-Biotech/ChemLib 143100 mg

From: Fredman, Jeffrey
Sent: Tuesday, January 25, 2005 6:05 AM
To: STIC-Biotech/ChemLib
Cc: Dunston, Jennifer
Subject: FW: Sequence Search 10/826523

PLEASE RUSH.

I Approve.

Jeff Fredman

-----Original Message-----
From: Dunston, Jennifer
Sent: Monday, January 24, 2005 12:46 PM
To: Fredman, Jeffrey
Subject: Sequence Search 10/826523

Jeff,
Please RUSH this search. The length of SEQ ID NO: 40 is 707 nucleotides. The length of SEQ ID NO: 41 is 3662 nucleotides. The sequences are related in that SEQ ID NO: 41 comprises the sequence of SEQ ID NO: 40.
Jenn

Please do a sequence search for the nucleic acid sequence of SEQ ID NO: 40 and SEQ ID NO: 41 against the commercial and interference nucleotide databases.
Thank you.

Jennifer Dunston, Ph.D.
USPTO Art Unit 1636
REM 2B76
Mailbox: REM 2C70
(571) 272-2916

****
STAFF USE ONLY
Searcher: Jan
Searcher Phone: 2-2504
Date Searcher Picked up: 1/25/05
Date Completed: 1/27/05
Searcher Prep/Rev. Time: 10
Online Time: 10

****
Type of Search
NA Sequence: # ✓
AA Sequence :#
Structure: #
Bibliographic:
Litigation:
Patent Family:
Other:

****
Vendors and cost where applicable
STN:
DIALOG:
QUESTEL/ORBIT:
LEXIS/NEXIS:
SEQUENCE SYSTEM: ✓
WWW/Internet:
Other(Specify):

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: January 25, 2005, 23:27:35 ; Search time 15598.5 Seconds
(without alignments)
11102.026 Million cell updates/sec

Title: US-10-826-523-41
Perfect score: 3662
Sequence: 1 ctaaattgtaagcgttaata.....atttccccgaaaagtgccac 3662

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Search Notes

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

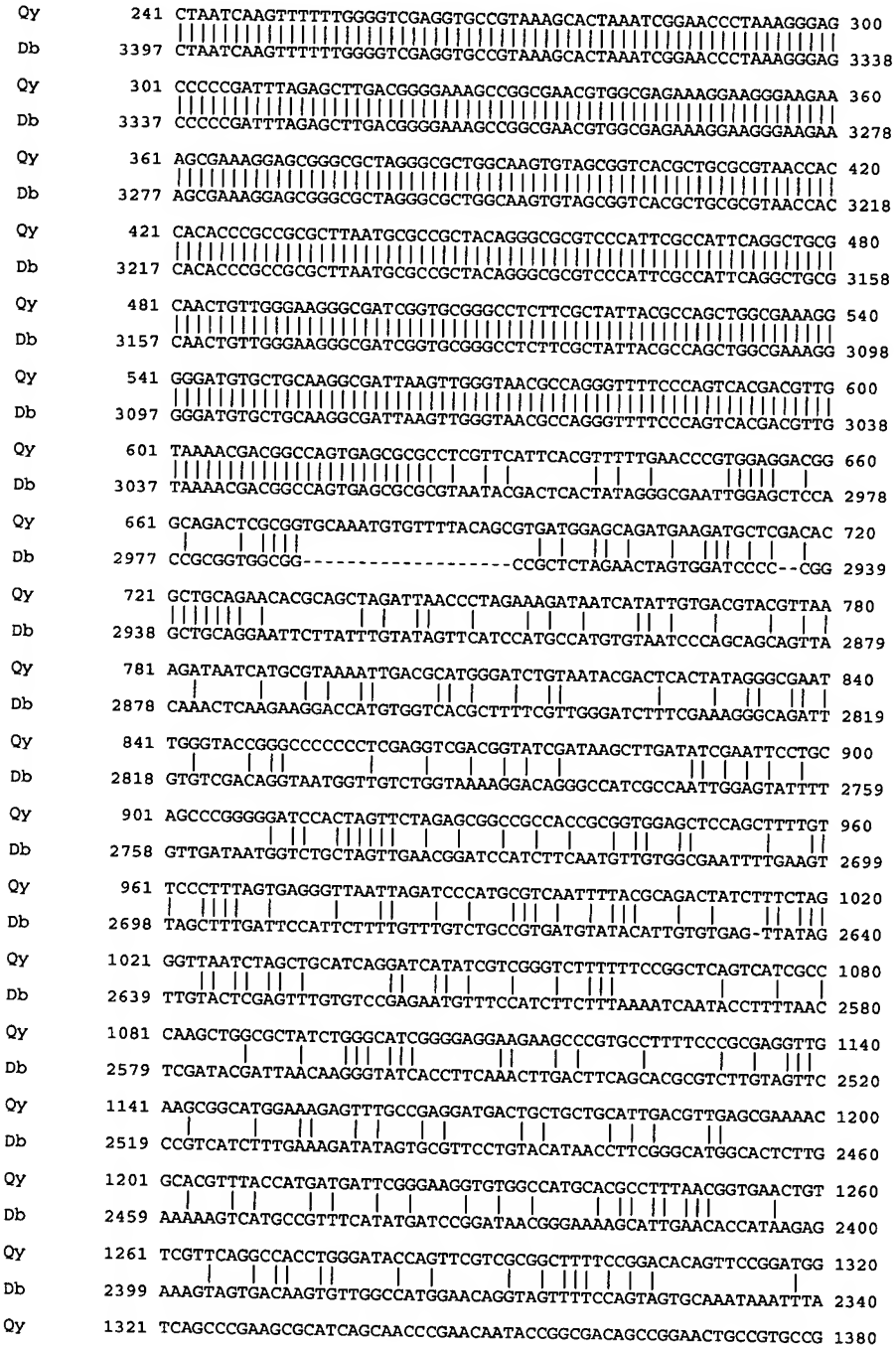
	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	2664.8	72.8	3637	6	AX592968	AX592968 Sequence
c	2	2656.4	72.5	3637	6	AX592967	AX592967 Sequence
c	3	2650.6	72.4	3626	12	ASPGREEN1	Y09373 Artificial
	4	2644.8	72.2	3624	12	XXU43957	U43957 Expression
	5	2639	72.1	3681	6	A20702	A20702 pKSEL5 DNA
	6	2639	72.1	3681	6	I15651	I15651 Sequence 6
	7	2623.6	71.6	5314	6	A20700	A20700 pMTV1 DNA s
	8	2623.6	71.6	5314	6	I15649	I15649 Sequence 4
c	9	2610.8	71.3	3552	12	AF171227	AF171227 Shuttle v
	10	2609.4	71.3	4001	6	AX306330	AX306330 Sequence
	11	2558.8	69.9	5540	12	AF179627	AF179627 Expressio
c	12	2550.6	69.7	3633	12	ASPGREEN2	Y09374 Artificial
c	13	2517.8	68.8	3918	6	CQ768840	CQ768840 Sequence
c	14	2517.8	68.8	3918	6	CQ802124	CQ802124 Sequence
	15	2514.6	68.7	3886	12	XXU35129	U35129 Plasmid pBS
c	16	2485.8	67.9	4141	12	AY190524	AY190524 Cloning v
	17	2484.6	67.8	3774	6	AX306326	AX306326 Sequence
	18	2474.2	67.6	5277	6	BD244458	BD244458 Genome an
	19	2469.4	67.4	5277	6	AR340565	AR340565 Sequence
	20	2451.8	67.0	3928	6	AX591027	AX591027 Sequence
	21	2451.8	67.0	3928	6	AX591180	AX591180 Sequence
	22	2451.8	67.0	3928	6	AX717603	AX717603 Sequence
	23	2449.8	66.9	7933	12	AF359244	AF359244 Integrati
	24	2448.8	66.9	3956	6	A84776	A84776 Sequence 21
	25	2448.8	66.9	3956	6	AR300179	AR300179 Sequence
	26	2448.8	66.9	3956	6	BD058066	BD058066 Method fo
	27	2446.8	66.8	3715	6	AX306328	AX306328 Sequence
	28	2434	66.5	3970	12	CVE403983	AJ403983 Cloning v
	29	2420.6	66.1	3274	12	XXU43956	U43956 Expression
c	30	2414.6	65.9	3978	12	CVU41513	U41513 3' terminal
c	31	2401	65.6	3733	12	AF146395	AF146395 Cloning v
	32	2396	65.4	3307	12	SYNPFABCX	D50399 Unidentifie
c	33	2382.6	65.1	4591	12	SYNPEZZ18A	M74186 Cloning vec
c	34	2380.4	65.0	6624	12	CVU37458	U37458 Yeast CUP1
	35	2372.2	64.8	4239	12	AVE401048	AJ401048 Artificia
	36	2359	64.4	5781	12	CVU37457	U37457 Yeast CUP1
	37	2357.4	64.4	4051	12	XXU35126	U35126 Plasmid pBS
c	38	2351	64.2	5275	12	CVU37456	U37456 Yeast CUP1
	39	2349.8	64.2	3230	12	AY663843	AY663843 Cloning v
	40	2340.2	63.9	3230	12	AY663844	AY663844 Cloning v
c	41	2333.4	63.7	5436	12	AY303166	AY303166 CFP Integ
c	42	2332.8	63.7	5436	12	AY303167	AY303167 YFP Integ
	43	2328.6	63.6	3204	12	AY663845	AY663845 Cloning v
c	44	2327.2	63.5	5385	12	IVE250107	AJ250107 Integrati
c	45	2322.6	63.4	5445	12	IVE306911	AJ306911 Integrati

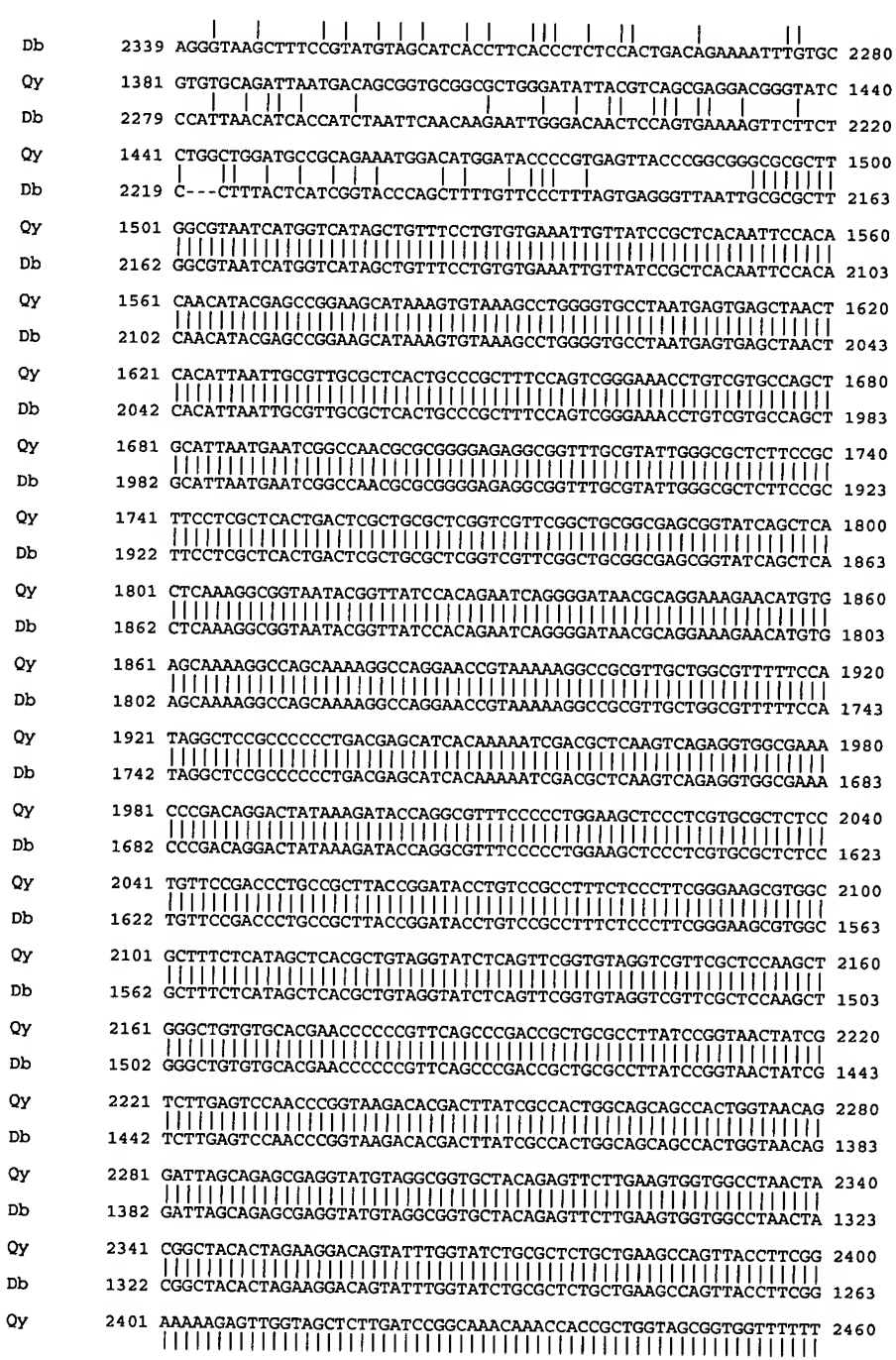
ALIGNMENTS

RESULT 1
AX592968/c
LOCUS AX592968 3637 bp DNA linear PAT 27-JAN-2003
DEFINITION Sequence 4 from Patent WO02079468.
ACCESSION AX592968
VERSION AX592968.1 GI:27950812
KEYWORDS .
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.
REFERENCE 1
AUTHORS Padgett,H.S., Fitzmaurice,W.P. and Lindo,J.A.
TITLE A method of increasing complementarity in a heteroduplex
JOURNAL Patent: WO 02079468-A 4 10-OCT-2002;
Large Scale Biology Corporation (US)
FEATURES Location/Qualifiers
source 1. .3637
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
ORIGIN

Query Match 72.8%; Score 2664.8; DB 6; Length 3637;
Best Local Similarity 83.8%; Pred. No. 0;
Matches 3070; Conservative 0; Mismatches 567; Indels 25; Gaps 4;

```
Qy          1 CTAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3637 CTAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTC 3578

Qy         61 ATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3577 ATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGA 3518

Qy        121 GATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3517 GATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTC 3458

Qy        181 CAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3457 CAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACC 3398
```





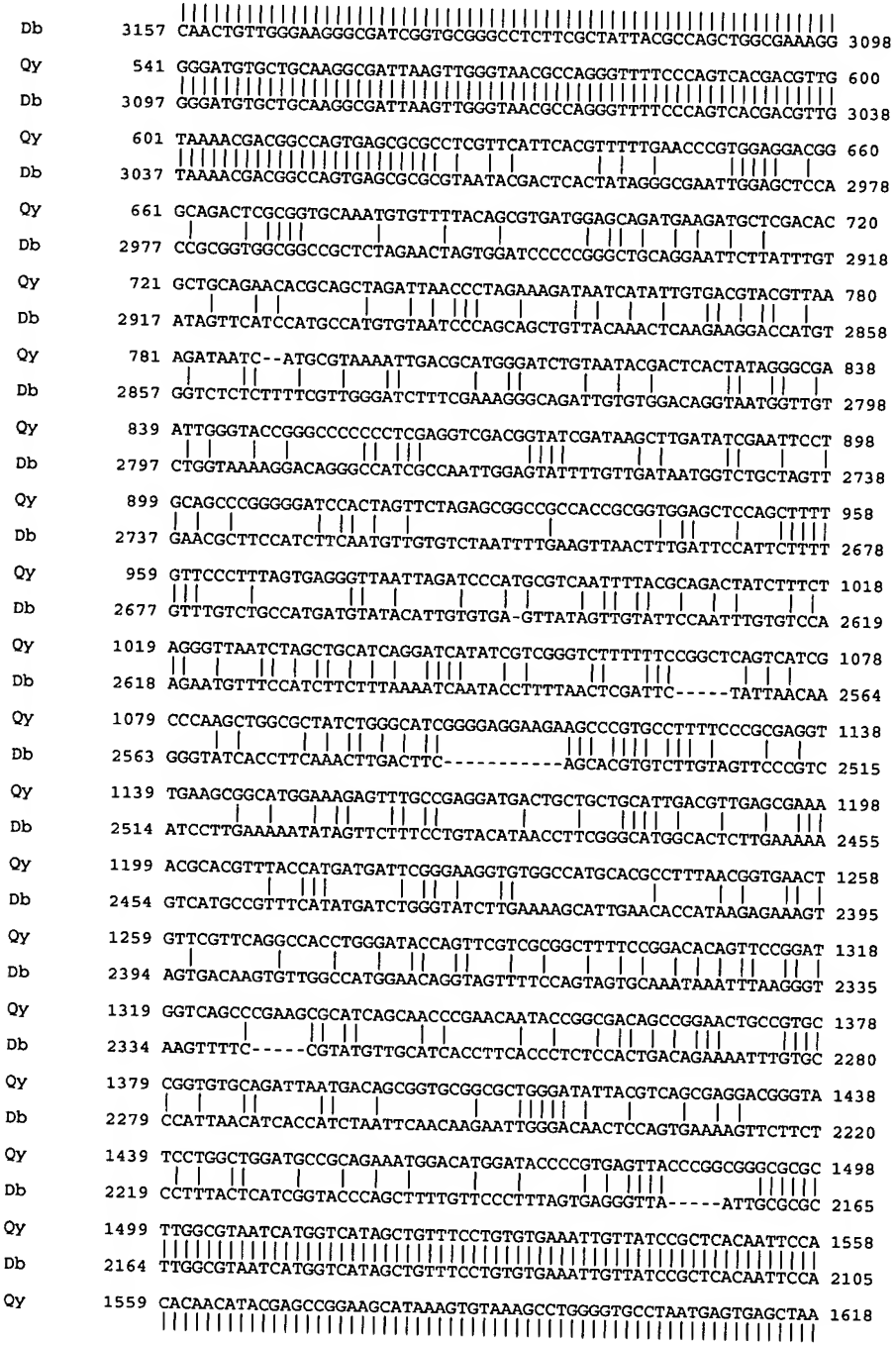
```
Db       1262 AAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTT 1203

Qy       2461 TGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTT 2520
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1202 TGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTT 1143

Qy       2521 TTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAG 2580
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1142 TTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAG 1083

Qy       2581 ATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAAT 2640
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1082 ATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAAT 1023

Qy       2641 CTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACC 2700
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1022 CTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACC 963

Qy       2701 TATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGAT 2760
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        962 TATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGAT 903

Qy       2761 AACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCC 2820
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        902 AACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCC 843

Qy       2821 ACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAG 2880
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        842 ACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAG 783

Qy       2881 AAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAG 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        782 AAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAG 723

Qy       2941 AGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGT 3000
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        722 AGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGT 663

Qy       3001 GGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCG 3060
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        662 GGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCG 603

Qy       3061 AGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGT 3120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        602 AGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGT 543

Qy       3121 TGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTC 3180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        542 TGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTC 483

Qy       3181 TCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTC 3240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        482 TCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTC 423

Qy       3241 ATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAA 3300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 ATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAA 363

Qy       3301 TACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCG 3360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 TACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCG 303

Qy       3361 AAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACC 3420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 AAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACC 243

Qy       3421 CAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAG 3480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 CAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAG 183

Qy       3481 GCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTT 3540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 GCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTT 123

Qy       3541 CCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATT 3600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         122 CCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATT 63

Qy       3601 TGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCC 3660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          62 TGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCC 3

Qy       3661 AC 3662
              ||
Db           2 AC 1
```

RESULT 2
AX592967/c

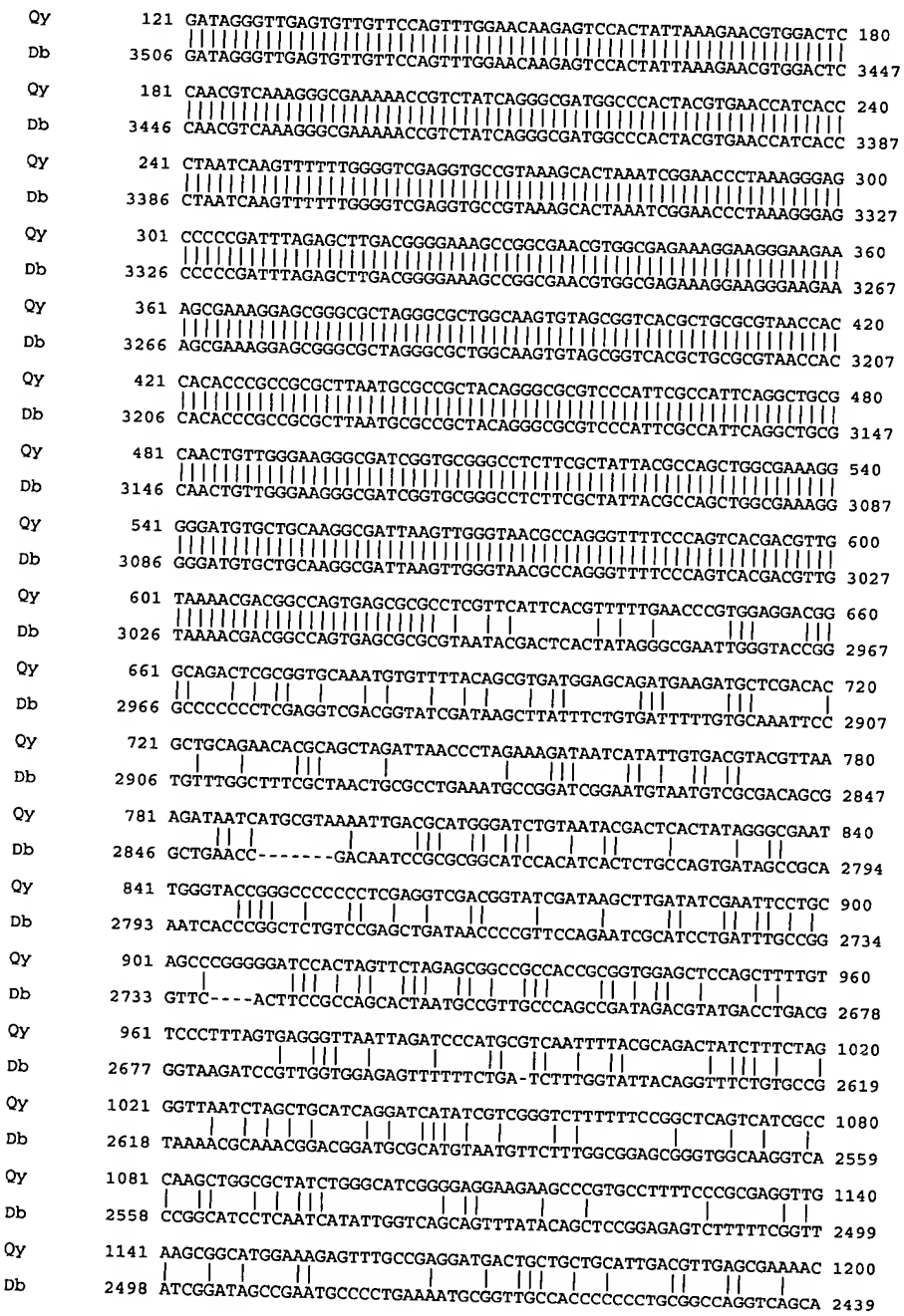
```
LOCUS       AX592967               3637 bp    DNA     linear   PAT 27-JAN-2003
DEFINITION  Sequence 3 from Patent WO02079468.
ACCESSION   AX592967
VERSION     AX592967.1  GI:27950811
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Padgett,H.S., Fitzmaurice,W.P. and Lindo,J.A.
  TITLE     A method of increasing complementarity in a heteroduplex
  JOURNAL   Patent: WO 02079468-A 3 10-OCT-2002;
            Large Scale Biology Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .3637
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
ORIGIN
```

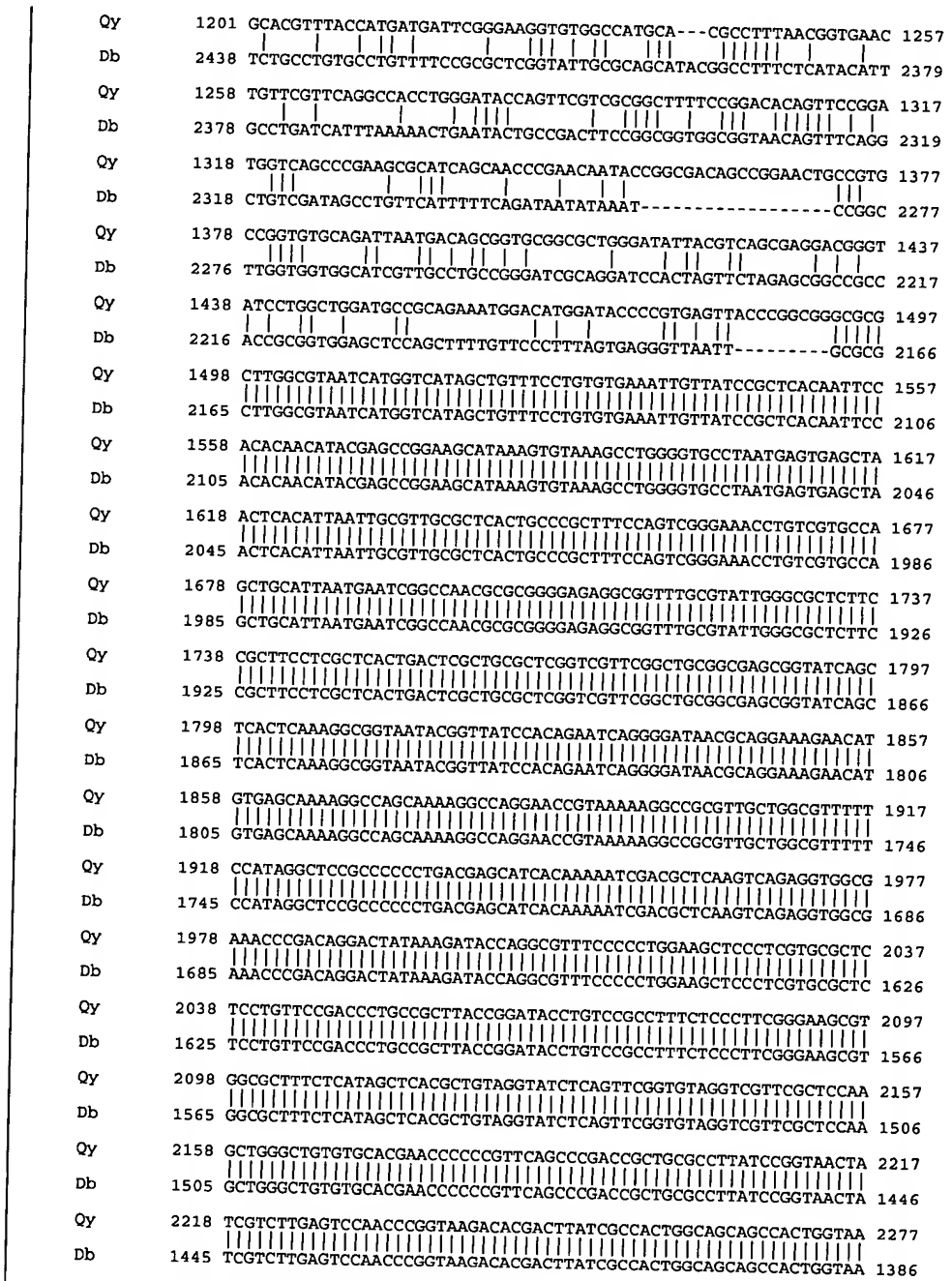
Query Match 72.5%; Score 2656.4; DB 6; Length 3637;
Best Local Similarity 84.0%; Pred. No. 0;
Matches 3079; Conservative 0; Mismatches 556; Indels 29; Gaps 6;

```
Qy          1 CTAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3637 CTAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTC 3578

Qy         61 ATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3577 ATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGA 3518

Qy        121 GATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3517 GATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTC 3458

Qy        181 CAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3457 CAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACC 3398

Qy        241 CTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3397 CTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAG 3338

Qy        301 CCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3337 CCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAA 3278

Qy        361 AGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCAC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3277 AGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCAC 3218

Qy        421 CACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCCCATTCGCCATTCAGGCTGCG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3217 CACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCCCATTCGCCATTCAGGCTGCG 3158

Qy        481 CAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGG 540
```


Db 3157 CAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGG 3098
Qy 541 GGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTG 600
Db 3097 GGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTG 3038
Qy 601 TAAAACGACGGCCAGTGAGCGCGCCTCGTTCATTCACGTTTTTGAACCCGTGGAGGACGG 660
Db 3037 TAAAACGACGGCCAGTGAGCGCGCGTAATACGACTCACTATAGGGCGAATTGGAGCTCCA 2978
Qy 661 GCAGACTCGCGGTGCAAATGTGTTTTACAGCGTGATGGAGCAGATGAAGATGCTCGACAC 720
Db 2977 CCGCGGTGGCGGCCGCTCTAGAACTAGTGGATCCCCCGGGCTGCAGGAATTCTTATTTGT 2918
Qy 721 GCTGCAGAACACGCAGCTAGATTAACCCTAGAAAGATAATCATATTGTGACGTACGTTAA 780
Db 2917 ATAGTTCATCCATGCCATGTGTAATCCCAGCAGCTGTTACAAACTCAAGAAGGACCATGT 2858
Qy 781 AGATAATC--ATGCGTAAAATTGACGCATGGGATCTGTAATACGACTCACTATAGGGCGA 838
Db 2857 GGTCTCTCTTTTCGTTGGGATCTTTCGAAAGGGCAGATTGTGTGGACAGGTAATGGTTGT 2798
Qy 839 ATTGGGTACCGGGCCCCCCCTCGAGGTCGACGGTATCGATAAGCTTGATATCGAATTCCT 898
Db 2797 CTGGTAAAAGGACAGGGCCATCGCCAATTGGAGTATTTTGTTGATAATGGTCTGCTAGTT 2738
Qy 899 GCAGCCCGGGGGATCCACTAGTTCTAGAGCGGCCGCCACCGCGGTGGAGCTCCAGCTTTT 958
Db 2737 GAACGCTTCCATCTTCAATGTTGTGTCTAATTTTGAAGTTAACTTTGATTCCATTCTTTT 2678
Qy 959 GTTCCCTTTAGTGAGGGTTAATTAGATCCCATGCGTCAATTTTACGCAGACTATCTTTCT 1018
Db 2677 GTTTGTCTGCCATGATGTATACATTGTGTGA-GTTATAGTTGTATTCCAATTTGTGTCCA 2619
Qy 1019 AGGGTTAATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCG 1078
Db 2618 AGAATGTTTCCATCTTCTTTAAAATCAATACCTTTTAACTCGATTC-----TATTAACAA 2564
Qy 1079 CCCAAGCTGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGT 1138
Db 2563 GGGTATCACCTTCAAACTTGACTTC----------AGCACGTGTCTTGTAGTTCCCGTC 2515
Qy 1139 TGAAGCGGCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAA 1198
Db 2514 ATCCTTGAAAAATATAGTTCTTTCCTGTACATAACCTTCGGGCATGGCACTCTTGAAAAA 2455
Qy 1199 ACGCACGTTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACT 1258
Db 2454 GTCATGCCGTTTCATATGATCTGGGTATCTTGAAAAGCATTGAACACCATAAGAGAAAGT 2395
Qy 1259 GTTCGTTCAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGAT 1318
Db 2394 AGTGACAAGTGTTGGCCATGGAACAGGTAGTTTTCCAGTAGTGCAAATAAATTTAAGGGT 2335
Qy 1319 GGTCAGCCCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGC 1378
Db 2334 AAGTTTTC-----CGTATGTTGCATCACCTTCACCCTCTCCACTGACAGAAAATTTGTGC 2280
Qy 1379 CGGTGTGCAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTA 1438
Db 2279 CCATTAACATCACCATCTAATTCAACAAGAATTGGGACAACTCCAGTGAAAAGTTCTTCT 2220
Qy 1439 TCCTGGCTGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGC 1498
Db 2219 CCTTTACTCATCGGTACCCAGCTTTTGTTCCCTTTAGTGAGGGTTA-----ATTGCGCGC 2165
Qy 1499 TTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCA 1558
Db 2164 TTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCA 2105
Qy 1559 CACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAA 1618
Db 2104 CACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAA 2045
Qy 1619 CTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAG 1678
Db 2044 CTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAG 1985
Qy 1679 CTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCC 1738
Db 1984 CTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCC 1925
Qy 1739 GCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCT 1798
Db 1924 GCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCT 1865
Qy 1799 CACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATG 1858
Db 1864 CACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATG 1805
Qy 1859 TGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTC 1918
Db 1804 TGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTC 1745
Qy 1919 CATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGA 1978
Db 1744 CATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGA 1685
Qy 1979 AACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCT 2038
Db 1684 AACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCT 1625
Qy 2039 CCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTG 2098
Db 1624 CCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTG 1565
Qy 2099 GCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAG 2158
Db 1564 GCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAG 1505
Qy 2159 CTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTAT 2218
Db 1504 CTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTAT 1445
Qy 2219 CGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAAC 2278
Db 1444 CGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAAC 1385
Qy 2279 AGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAAC 2338
Db 1384 AGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAAC 1325
Qy 2339 TACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTC 2398
Db 1324 TACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTC 1265
Qy 2399 GGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTT 2458
Db 1264 GGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTT 1205
Qy 2459 TTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATC 2518
Db 1204 TTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATC 1145
Qy 2519 TTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATG 2578
Db 1144 TTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATG 1085
Qy 2579 AGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCA 2638
Db 1084 AGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCA 1025
Qy 2639 ATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCA 2698
Db 1024 ATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCA 965

```
Qy       2699 CCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAG 2758
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        964 CCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAG 905

Qy       2759 ATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGAC 2818
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        904 ATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGAC 845

Qy       2819 CCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGC 2878
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        844 CCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGC 785

Qy       2879 AGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCT 2938
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        784 AGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCT 725

Qy       2939 AGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATC 2998
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        724 AGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATC 665

Qy       2999 GTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGG 3058
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        664 GTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGG 605

Qy       3059 CGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATC 3118
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        604 CGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATC 545

Qy       3119 GTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAAT 3178
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        544 GTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAAT 485

Qy       3179 TCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAG 3238
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        484 TCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAG 425

Qy       3239 TCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGAT 3298
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        424 TCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGAT 365

Qy       3299 AATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGG 3358
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        364 AATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGG 305

Qy       3359 CGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCA 3418
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        304 CGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCA 245

Qy       3419 CCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGA 3478
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 CCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGA 185

Qy       3479 AGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTC 3538
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        184 AGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTC 125

Qy       3539 TTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATA 3598
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        124 TTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATA 65

Qy       3599 TTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTG 3658
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          64 TTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTG 5

Qy       3659 CCAC 3662
              ||||
Db          4 CCAC 1


RESULT 3
ASPGREEN1/c
LOCUS       ASPGREEN1               3626 bp    DNA     linear   SYN 15-MAY-1998
DEFINITION  Artificial sequences, plasmid vector pGreen-1.
ACCESSION   Y09373
VERSION     Y09373.1  GI:1684624
KEYWORDS    blaTEM gene; lacZ-PhoC fusion gene; multiple cloning site; T3
            promoter.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Thaller,M.C., Berlutti,F., Schippa,S., Selan,L. and Rossolini,G.M.
  TITLE     Bacterial acid phosphatase gene fusions useful as targets for
            cloning-dependent insertional inactivation
  JOURNAL   Biotechnol. Prog. 14 (2), 241-247 (1998)
  MEDLINE   98215688
   PUBMED   9548775
REFERENCE   2  (bases 1 to 3626)
  AUTHORS   Rossolini,G.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-NOV-1996) G.M. Rossolini, Dip. Biologia
            Moleculare-Sez., Microbiologia,, Univ. di Siena, via Laterina N.8,
            I- 53100 Siena, ITALY
COMMENT     Related sequences: X64444 and X52328.
FEATURES             Location/Qualifiers
     source          1. .3626
                     /organism="synthetic construct"
                     /mol_type="other DNA"
                     /db_xref="taxon:32630"
                     /plasmid="pGreen-1"
     gene            129. .989
                     /gene="blaTEM"
     CDS             129. .989
                     /gene="blaTEM"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAA70547.1"
                     /db_xref="GI:1684625"
                     /translation="MSIQHFRVALIPFFAAFCLPVFAHPETLVKVKDAEDQLGARVGY
                     IELDLNSGKILESFRPEERFPMMSTFKVLLCGAVLSRIDAGQEQLGRRIHYSQNDLVE
                     YSPVTEKHLTDGMTVRELCSAAITMSDNTAANLLLTTIGGPKELTAFLHNMGDHVTRL
                     DRWEPELNEAIPNDERDTTMPVAMATTLRKLLTGELLTLASRQQLIDWMEADKVAGPL
                     LRSALPAGWFIADKSGAGERGSRGIIAALGPDGKPSRIVVIYTTGSQATMDERNRQIA
                     EIGASLIKHW"
     primer_bind     2139. .2154
                     /note="pUC M13/Reverse Sequencing primer recognition site"
     gene            2146. .2934
                     /gene="lacZ-PhoC fusion"
     CDS             2146. .2934
                     /gene="lacZ-PhoC fusion"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAA70548.1"
                     /db_xref="GI:1684626"
                     /translation="MTMITPSAQLTLTKGNKSWSSTAVAAALELVDPAIPAGNDATTK
                     PDLYYLKNEQAIDSLKLLPPPPEVGSIQFLNDQAMYEKGRMLRNTERGKQAQADADLA
                     AGGVATAFSGAFGYPITEKDSPELYKLLTNMIEDAGDLATRSAKEHYMRIRPFAFYGT
                     ETCNTKDQKKLSTNGSYPSGHTSIGWATALVLAEVNPANQDAILERGYQLGQSRVICG
                     YHWQSDVDAARIVGSAAVATLHSDPAFQAQLAKAKQEFAQKSQK"
     promoter        2171. .2190
                     /gene="lacZ-PhoC fusion"
                     /note="T3 promoter"
ORIGIN

  Query Match           72.4%;  Score 2650.6;  DB 12;  Length 3626;
  Best Local Similarity   84.0%;  Pred. No. 0;
  Matches 3079;  Conservative    0;  Mismatches  544;  Indels   42;  Gaps    6;

Qy          1 CTAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3626 CTAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTC 3567

Qy         61 ATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3566 ATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGA 3507
```





```
Db       1385 CAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAA 1326

Qy       2338 CTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTT 2397
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1325 CTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTT 1266

Qy       2398 CGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTT 2457
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1265 CGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTT 1206

Qy       2458 TTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGAT 2517
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1205 TTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGAT 1146

Qy       2518 CTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCAT 2577
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1145 CTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCAT 1086

Qy       2578 GAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATC 2637
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1085 GAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATC 1026

Qy       2638 AATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGC 2697
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1025 AATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGC 966

Qy       2698 ACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTA 2757
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        965 ACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTA 906

Qy       2758 GATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGA 2817
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        905 GATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGA 846

Qy       2818 CCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCG 2877
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        845 CCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCG 786

Qy       2878 CAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGC 2937
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        785 CAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGC 726

Qy       2938 TAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCAT 2997
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        725 TAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCAT 666

Qy       2998 CGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAG 3057
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        665 CGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAG 606

Qy       3058 GCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGAT 3117
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        605 GCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGAT 546

Qy       3118 CGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAA 3177
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        545 CGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAA 486

Qy       3178 TTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAA 3237
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        485 TTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAA 426

Qy       3238 GTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGA 3297
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        425 GTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGA 366

Qy       3298 TAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGG 3357
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        365 TAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGG 306

Qy       3358 GCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGC 3417
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        305 GCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGC 246

Qy       3418 ACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGG 3477
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        245 ACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGG 186

Qy       3478 AAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACT 3537
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        185 AAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACT 126

Qy       3538 CTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACAT 3597
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        125 CTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACAT 66

Qy       3598 ATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGT 3657
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         65 ATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGT 6

Qy       3658 GCCAC 3662
              |||||
Db          5 GCCAC 1
```

RESULT 4
XXU43957

```
LOCUS       XXU43957               3624 bp    DNA     linear   SYN 28-JAN-1999
DEFINITION  Expression vector pB10T-MRz, complete cds.
ACCESSION   U43957
VERSION     U43957.1  GI:4097019
KEYWORDS    .
SOURCE      Expression vector pB10T-MRz
  ORGANISM  Expression vector pB10T-MRz
            artificial sequences; vectors.
REFERENCE   1  (bases 1 to 3624)
  AUTHORS   Ruiz,J., Wu,C.H., Ito,Y. and Wu,G.Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-DEC-1995) Juan Ruiz, Division of
            Gastroenterology-Hepatology, University of Connecticut Health
            Center, Farmington, CT 06030, USA
COMMENT     In vivo expression vector for the synthesis of a ribozyme against
            the core region of the Hepatitis B virus pregenomic RNA. It does
            not recognize the viral core protein but the core region of the
            viral RNA. pB10T-MRz contains a tandem of ten units of the
            ribozyme, separated between them by the target sequence recognized
            by the ribozyme, cloned at the XbaI-HindIII site of pBluescript SK
            (+).
FEATURES             Location/Qualifiers
     source          1. .3624
                     /organism="Expression vector pB10T-MRz"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:45892"
     misc_feature    complement(657. .1425)
                     /note="multiple cloning site"
     misc_RNA        complement(721. .766)
                     /product="hammerhead ribozyme"
     misc_RNA        complement(791. .836)
                     /product="hammerhead ribozyme"
     misc_RNA        complement(861. .906)
                     /product="hammerhead ribozyme"
     misc_RNA        complement(931. .976)
                     /product="hammerhead ribozyme"
     misc_RNA        complement(1001. .1046)
                     /product="hammerhead ribozyme"
     misc_RNA        complement(1071. .1116)
                     /product="hammerhead ribozyme"
     misc_RNA        complement(1141. .1186)
                     /product="hammerhead ribozyme"
     misc_RNA        complement(1211. .1256)
                     /product="hammerhead ribozyme"
     misc_RNA        complement(1281. .1326)
                     /product="hammerhead ribozyme"
     misc_RNA        complement(1351. .1396)
                     /product="hammerhead ribozyme"
```

```
     promoter        complement(1482. .1604)
                     /note="Lac-Z"
     gene            complement(2639. .3499)
                     /gene="ampR"
     CDS             complement(2639. .3499)
                     /gene="ampR"
                     /function="ampicillin resistance marker"
                     /codon_start=1
                     /transl_table=11
                     /product="beta-lactamase"
                     /protein_id="AAD10137.1"
                     /db_xref="GI:4097020"
                     /translation="MSIQHFRVALIPFFAAFCLPVFAHPETLVKVKDAEDQLGARVGY
                     IELDLNSGKILESFRPEERFPMMSTFKVLLCGAVLSRIDAGQEQLGRRIHYSQNDLVE
                     YSPVTEKHLTDGMTVRELCSAAITMSDNTAANLLLTTIGGPKELTAFLHNMGDHVTRL
                     DRWEPELNEAIPNDERDTTMPVAMATTLRKLLTGELLTLASRQQLIDWMEADKVAGPL
                     LRSALPAGWFIADKSGAGERGSRGIIAALGPDGKPSRIVVIYTTGSQATMDERNRQIA
                     EIGASLIKHW"
ORIGIN

  Query Match          72.2%;  Score 2644.8;  DB 12;  Length 3624;
  Best Local Similarity  84.3%;  Pred. No. 0;
  Matches 3096;  Conservative    0;  Mismatches  512;  Indels   64;  Gaps    8;

Qy        1 CTAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTC 60
Db        4 CTAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTC 63

Qy       61 ATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGA 120
Db       64 ATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGA 123

Qy      121 GATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTC 180
Db      124 GATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTC 183

Qy      181 CAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACC 240
Db      184 CAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACC 243

Qy      241 CTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAG 300
Db      244 CTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAG 303

Qy      301 CCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAA 360
Db      304 CCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAA 363

Qy      361 AGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCAC 420
Db      364 AGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCAC 423

Qy      421 CACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCCCATTCGCCATTCAGGCTGCG 480
Db      424 CACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCCCATTCGCCATTCAGGCTGCG 483

Qy      481 CAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGG 540
Db      484 CAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGG 543

Qy      541 GGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTG 600
Db      544 GGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTG 603

Qy      601 TAAAACGACGGCCAGTG------AGCGCGCCTCGTTCATTCACGTTTTTGAACCCGTGGA 654
Db      604 TAAAACGACGGCCAGTGAATTGTAATACGACTCACTATAGGGCGAATTGGGTACCGGGCC 663

Qy      655 GGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGCGTGATGGAGCAGATGAAGATGCT 714
Db      664 CCCCCTCGAGGTCGACGGTATCGATAAGCTTGGACTAGTTCAGGAGACTCTGGATCCAAG 723

Qy      715 CGACACGCTGCAGAACACGCAGCTAGATTAACCCTAGAAAGATAATCATATTGTGACGTA 774

Db      724 CCTTAGAGTTTCGTCCTCACGGACTCATCAGTCCT---GAGCATTGTCTAGTTCAGGAGA 780

Qy      775 CGTTAAAGATAATCATGCGTAAAATTGACGCATGGGATCTGTAATACGACTCACTATAGG 834

Db      781 CTCTGGATCCAAGCCTTAGAGTTTCGTCCTCACGGACTCATCAGTCCTGAGCATTGTCTA 840

Qy      835 GCGAATTGGGTACCGGGCCCCCCCTCGAGGTCGACGGTATCGATAAGCTTGATATCGAAT 894

Db      841 GTTCAGGAGACTCTGGATCCAAGCCTTAGAGTTTCGTCCTCACGGA---------------- 886

Qy      895 TCCTGCAGCCCGGGGGATCCACTAGTTCTAGAGCGGCCGCCACCGCGGTGGAGCTCCAGC 954

Db      887 CTCATCAGTCCTGAGCATTGTCTAGTTCAGGAGA------------------------C 921

Qy      955 TTTTGTTCCCTTTAGTGAGGGTTAATTAGATCCCATGCGTCAATTTTACGCAGACTATCT 1014

Db      922 TCTGGATCCAAGCCTTAGAGTTTCGTCCTCACGGACTCATCAGTCCTGAGCATTGTCTAG 981

Qy     1015 TTCTAGGGTTAATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTC 1074

Db      982 TTCAGGAG------ACTCTGGATCCAAGCCTTAGAGTTTCGTC--CTCACGGACTCATCA 1033

Qy     1075 ATCGCCCAAGCTGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCG 1134

Db     1034 GTCCTGAGCATTGTCTAGTTCAGGAGACTCTGGATCCAAGCCTTAGAGTTTCGTCCTCAC 1093

Qy     1135 AGGTTGAAGCGGCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGC 1194

Db     1094 GGACTCATCAGTCCTGAGCATTGTCTAGTTCAGGAGACT-CTGGATCCAAGCCTTAGAGT 1152

Qy     1195 GAAAACGCACGTTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTG 1254

Db     1153 TTCGTCCTCACGGACTCATCAGTCCTGAGCATTGTCTAGTTCAGGAGACTCTGGATCCAA 1212

Qy     1255 AACTGTTCGTTCAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCC 1314

Db     1213 GCCTTAGAGTTTCGTCCTCACGGACTCATCAGTCCTGAGCATTGTCTAGTTCAGGAGACT 1272

Qy     1315 GGATGGTCAGCCCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCC 1374

Db     1273 CTGGATCCAAGCCTTAGAGTTTCGTCCTCACGGACTCATCAGTCCTGAGCATTGTCTAGT 1332

Qy     1375 GTGCCGGTGTGCAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACG 1434

Db     1333 TCAGGAGACTCTGGATCCAAGCCTTAGAGTTTCGTCCTCACGGACTCATCAGTCCTGAGC 1392

Qy     1435 GGTATCCTGGCTGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGG- 1493

Db     1393 ATTGTCTAGAGCGGCCGCCACCGCGGTGGAGCTCCAGCTTTTGTTCCCTTTAGTGAGGGT 1452

Qy     1494 ------CGCGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGC 1547

Db     1453 TAATTTCGAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGC 1512

Qy     1548 TCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAAT 1607

Db     1513 TCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAAT 1572

Qy     1608 GAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACC 1667

Db     1573 GAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACC 1632

Qy     1668 TGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTG 1727

Db     1633 TGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTG 1692

Qy     1728 GGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAG 1787

Db     1693 GGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAG 1752

Qy     1788 CGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAG 1847
```

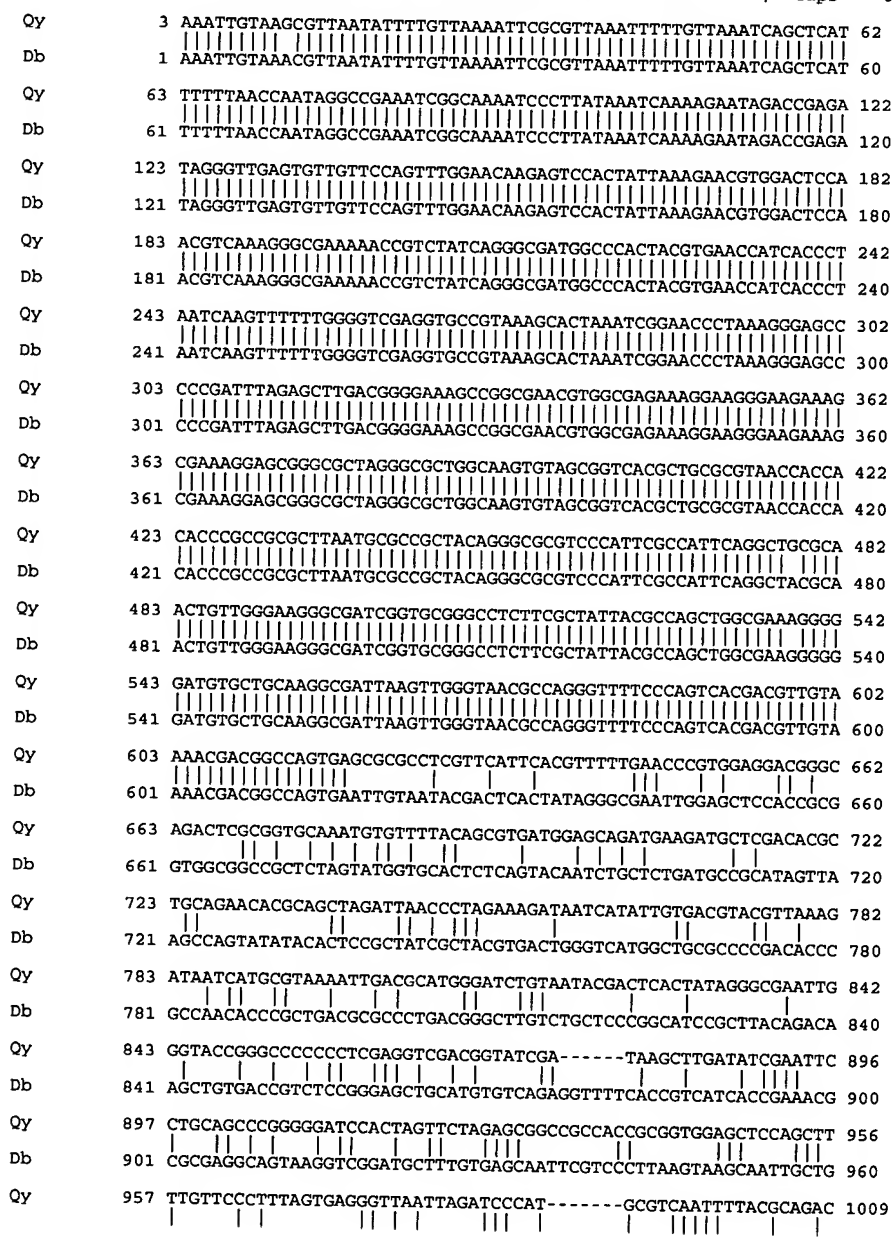
```
Db       1753 CGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAG 1812

Qy       1848 GAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGC 1907
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1813 GAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGC 1872

Qy       1908 TGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTC 1967
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1873 TGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTC 1932

Qy       1968 AGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCC 2027
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1933 AGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCC 1992

Qy       2028 TCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTT 2087
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1993 TCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTT 2052

Qy       2088 CGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCG 2147
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2053 CGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCG 2112

Qy       2148 TTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTAT 2207
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2113 TTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTAT 2172

Qy       2208 CCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAG 2267
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2173 CCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAG 2232

Qy       2268 CCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGT 2327
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2233 CCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGT 2292

Qy       2328 GGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGC 2387
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2293 GGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGC 2352

Qy       2388 CAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTA 2447
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2353 CAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTA 2412

Qy       2448 GCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAG 2507
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2413 GCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAG 2472

Qy       2508 ATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGA 2567
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2473 ATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGA 2532

Qy       2568 TTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAA 2627
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2533 TTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAA 2592

Qy       2628 GTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAA 2687
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2593 GTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAA 2652

Qy       2688 TCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCC 2747
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2653 TCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCC 2712

Qy       2748 CCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGA 2807
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2713 CCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGA 2772

Qy       2808 TACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAA 2867
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2773 TACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAA 2832

Qy       2868 GGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTT 2927
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2833 GGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTT 2892

Qy       2928 GCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTG 2987
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2893 GCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTG 2952

Qy       2988 CTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCC 3047
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2953 CTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCC 3012

Qy       3048 AACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCG 3107
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3013 AACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCG 3072

Qy       3108 GTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAG 3167
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3073 GTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAG 3132

Qy       3168 CACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGT 3227
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3133 CACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGT 3192

Qy       3228 ACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGT 3287
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3193 ACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGT 3252

Qy       3288 CAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAAC 3347
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3253 CAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAAC 3312

Qy       3348 GTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAAC 3407
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3313 GTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAAC 3372

Qy       3408 CCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAG 3467
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3373 CCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAG 3432

Qy       3468 CAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAA 3527
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3433 CAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAA 3492

Qy       3528 TACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGA 3587
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3493 TACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGA 3552

Qy       3588 GCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTC 3647
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3553 GCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTC 3612

Qy       3648 CCCGAAAAGTGC 3659
             ||||||||||||
Db       3613 CCCGAAAAGTGC 3624
```

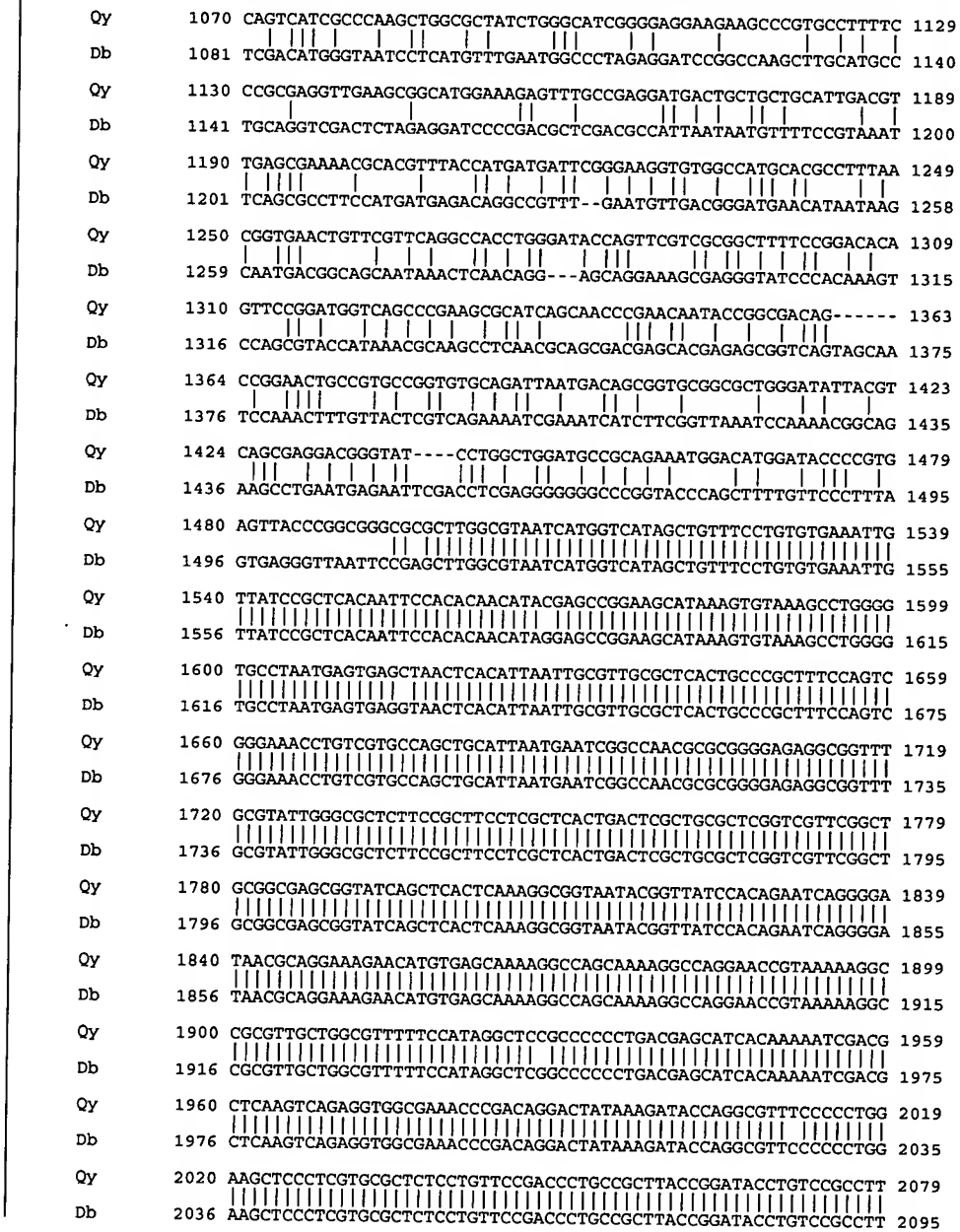
```
RESULT 5
A20702
LOCUS       A20702                 3681 bp    DNA     linear   PAT 22-AUG-1994
DEFINITION  pKSEL5 DNA sequence.
ACCESSION   A20702
VERSION     A20702.1  GI:583404
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 3681)
  AUTHORS   .
  TITLE     CARRIER-BOUND RECOMBINANT PROTEIN, PROCESS FOR PRODUCING IT AND ITS
            USE AS AN IMMUNOGEN AND VACCINE
  JOURNAL   Patent: WO 9113155-A 6 05-SEP-1991;
FEATURES             Location/Qualifiers
     source          1. .3681
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
```

/db_xref="taxon:32630"

ORIGIN

Query Match 72.1%; Score 2639; DB 6; Length 3681;
Best Local Similarity 83.5%; Pred. No. 0;
Matches 3075; Conservative 0; Mismatches 580; Indels 28; Gaps 6;





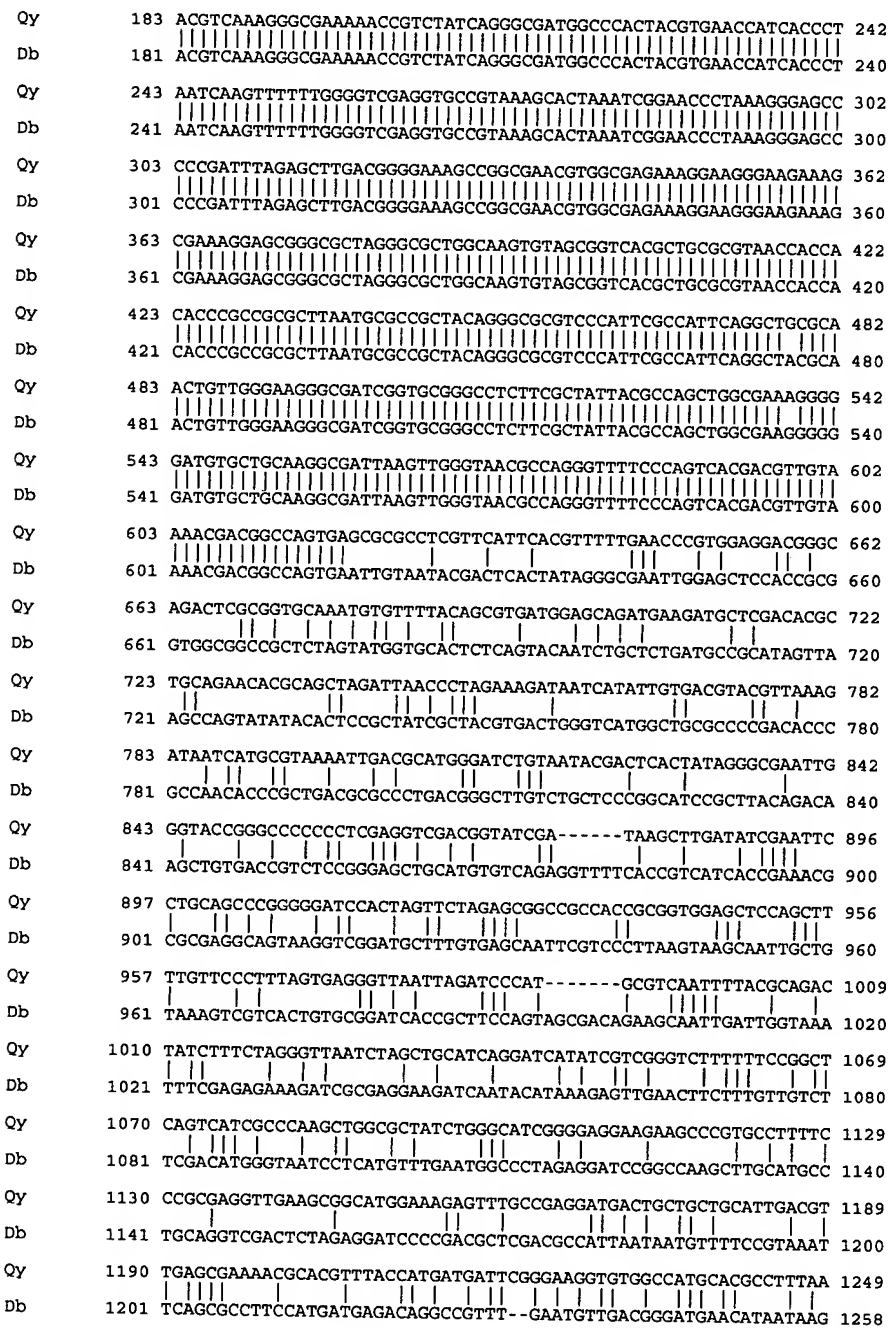
```
Qy       2080 TCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGT 2139
             ||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||||
Db       2096 TCTCCCTTCGGGAAGCGTGGCGCTTTCTCAATGCTCACGCTGTAGGTATCTCAGTTCGGT 2155

Qy       2140 GTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTG 2199
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2156 GTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTG 2215

Qy       2200 CGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACT 2259
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2216 CGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACT 2275

Qy       2260 GGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTT 2319
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2276 GGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTT 2335

Qy       2320 CTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCT 2379
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2336 CTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCT 2395

Qy       2380 GCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCAC 2439
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2396 GCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCAC 2455

Qy       2440 CGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATC 2499
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2456 CGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATC 2515

Qy       2500 TCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACG 2559
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2516 TCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACG 2575

Qy       2560 TTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTA 2619
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2576 TTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTA 2635

Qy       2620 AAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCA 2679
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2636 AAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCA 2695

Qy       2680 ATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGC 2739
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2696 ATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGC 2755

Qy       2740 CTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGC 2799
             |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2756 CTGACTGCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGC 2815

Qy       2800 TGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCC 2859
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2816 TGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCC 2875

Qy       2860 AGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTAT 2919
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2876 AGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTAT 2935

Qy       2920 TAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGT 2979
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2936 TAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGT 2995

Qy       2980 TGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTC 3039
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2996 TGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTC 3055

Qy       3040 CGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAG 3099
             |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db       3056 CGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGAAAAAAAGCGGTTAG 3115

Qy       3100 CTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGT 3159
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db       3116 CTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGCT 3175

Qy       3160 TATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGAC 3219
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3176 TATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGAC 3235

Qy       3220 TGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTG 3279
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3236 TGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTG 3295

Qy       3280 CCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCAT 3339
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3296 CCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCAT 3355

Qy       3340 TGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTC 3399
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3356 TGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTC 3415

Qy       3400 GATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTC 3459
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3416 GATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTC 3475

Qy       3460 TGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAA 3519
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3476 TGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAA 3535

Qy       3520 ATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTG 3579
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3536 ATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTG 3595

Qy       3580 TCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCG 3639
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3596 TCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCG 3655

Qy       3640 CACATTTCCCCGAAAAGTGCCAC 3662
             |||||||||||||||||||||||
Db       3656 CACATTTCCCCGAAAAGTGCCAC 3678
```

RESULT 6
I15651

```
LOCUS       I15651                 3681 bp    DNA     linear   PAT 02-APR-1996
DEFINITION  Sequence 6 from patent US 5470573.
ACCESSION   I15651
VERSION     I15651.1  GI:1250559
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 3681)
  AUTHORS   Lubitz,W. and Szostak,M.P.
  TITLE     Immunogens comprising the non-lytic membrane spanning domain of
            bacteriophages MS2 or PhiX174
  JOURNAL   Patent: US 5470573-A 6 28-NOV-1995;
FEATURES             Location/Qualifiers
     source          1. .3681
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN
```

Query Match 72.1%; Score 2639; DB 6; Length 3681;
Best Local Similarity 83.5%; Pred. No. 0;
Matches 3075; Conservative 0; Mismatches 580; Indels 28; Gaps 6;

```
Qy          3 AAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTCAT 62
              ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AAATTGTAAACGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTCAT 60

Qy         63 TTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGAGA 122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGAGA 120

Qy        123 TAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTCCA 182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTCCA 180
```


Qy 183 ACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCCT 242
Db 181 ACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCCT 240
Qy 243 AATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAGCC 302
Db 241 AATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAGCC 300
Qy 303 CCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAAAG 362
Db 301 CCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAAAG 360
Qy 363 CGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCACCA 422
Db 361 CGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCACCA 420
Qy 423 CACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCCCATTCGCCATTCAGGCTGCGCA 482
Db 421 CACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCCCATTCGCCATTCAGGCTACGCA 480
Qy 483 ACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGGG 542
Db 481 ACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAGGGGG 540
Qy 543 GATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTGTA 602
Db 541 GATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTGTA 600
Qy 603 AAACGACGGCCAGTGAGCGCGCCTCGTTCATTCACGTTTTTGAACCCGTGGAGGACGGGC 662
Db 601 AAACGACGGCCAGTGAATTGTAATACGACTCACTATAGGGCGAATTGGAGCTCCACCGCG 660
Qy 663 AGACTCGCGGTGCAAATGTGTTTTACAGCGTGATGGAGCAGATGAAGATGCTCGACACGC 722
Db 661 GTGGCGGCCGCTCTAGTATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTA 720
Qy 723 TGCAGAACACGCAGCTAGATTAACCCTAGAAAGATAATCATATTGTGACGTACGTTAAAG 782
Db 721 AGCCAGTATATACACTCCGCTATCGCTACGTGACTGGGTCATGGCTGCGCCCCGACACCC 780
Qy 783 ATAATCATGCGTAAAATTGACGCATGGGATCTGTAATACGACTCACTATAGGGCGAATTG 842
Db 781 GCCAACACCCGCTGACGCGCCCTGACGGGCTTGTCTGCTCCCGGCATCCGCTTACAGACA 840
Qy 843 GGTACCGGGCCCCCCCTCGAGGTCGACGGTATCGA------TAAGCTTGATATCGAATTC 896
Db 841 AGCTGTGACCGTCTCCGGGAGCTGCATGTGTCAGAGGTTTTCACCGTCATCACCGAAACG 900
Qy 897 CTGCAGCCCGGGGGATCCACTAGTTCTAGAGCGGCCGCCACCGCGGTGGAGCTCCAGCTT 956
Db 901 CGCGAGGCAGTAAGGTCGGATGCTTTGTGAGCAATTCGTCCCTTAAGTAAGCAATTGCTG 960
Qy 957 TTGTTCCCTTTAGTGAGGGTTAATTAGATCCCAT-------GCGTCAATTTTACGCAGAC 1009
Db 961 TAAAGTCGTCACTGTGCGGATCACCGCTTCCAGTAGCGACAGAAGCAATTGATTGGTAAA 1020
Qy 1010 TATCTTTCTAGGGTTAATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCT 1069
Db 1021 TTTCGAGAGAAAGATCGCGAGGAAGATCAATACATAAAGAGTTGAACTTCTTTGTTGTCT 1080
Qy 1070 CAGTCATCGCCCAAGCTGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTC 1129
Db 1081 TCGACATGGGTAATCCTCATGTTTGAATGGCCCTAGAGGATCCGGCCAAGCTTGCATGCC 1140
Qy 1130 CCGCGAGGTTGAAGCGGCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGT 1189
Db 1141 TGCAGGTCGACTCTAGAGGATCCCCGACGCTCGACGCCATTAATAATGTTTTCCGTAAAT 1200
Qy 1190 TGAGCGAAAACGCACGTTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAA 1249
Db 1201 TCAGCGCCTTCCATGATGAGACAGGCCGTTT--GAATGTTGACGGGATGAACATAATAAG 1258
Qy 1250 CGGTGAACTGTTCGTTCAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACA 1309
Db 1259 CAATGACGGCAGCAATAAACTCAACAGG---AGCAGGAAAGCGAGGGTATCCCACAAAGT 1315
Qy 1310 GTTCCGGATGGTCAGCCCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAG------ 1363
Db 1316 CCAGCGTACCATAAACGCAAGCCTCAACGCAGCGACGAGCACGAGAGCGGTCAGTAGCAA 1375
Qy 1364 CCGGAACTGCCGTGCCGGTGTGCAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGT 1423
Db 1376 TCCAAACTTTGTTACTCGTCAGAAAATCGAAATCATCTTCGGTTAAATCCAAAACGGCAG 1435
Qy 1424 CAGCGAGGACGGGTAT----CCTGGCTGGATGCCGCAGAAATGGACATGGATACCCCGTG 1479
Db 1436 AAGCCTGAATGAGAATTCGACCTCGAGGGGGGGCCCGGTACCCAGCTTTTGTTCCCTTTA 1495
Qy 1480 AGTTACCCGGCGGGCGCGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTG 1539
Db 1496 GTGAGGGTTAATTCCGAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTG 1555
Qy 1540 TTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGG 1599
Db 1556 TTATCCGCTCACAATTCCACACAACATAGGAGCCGGAAGCATAAAGTGTAAAGCCTGGGG 1615
Qy 1600 TGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTC 1659
Db 1616 TGCCTAATGAGTGAGGTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTC 1675
Qy 1660 GGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTT 1719
Db 1676 GGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTT 1735
Qy 1720 GCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCT 1779
Db 1736 GCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCT 1795
Qy 1780 GCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGA 1839
Db 1796 GCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGA 1855
Qy 1840 TAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGC 1899
Db 1856 TAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGC 1915
Qy 1900 CGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACG 1959
Db 1916 CGCGTTGCTGGCGTTTTTCCATAGGCTCGGCCCCCCTGACGAGCATCACAAAAATCGACG 1975
Qy 1960 CTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGG 2019
Db 1976 CTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTCCCCCCTGG 2035
Qy 2020 AAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTT 2079
Db 2036 AAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTT 2095
Qy 2080 TCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGT 2139
Db 2096 TCTCCCTTCGGGAAGCGTGGCGCTTTCTCAATGCTCACGCTGTAGGTATCTCAGTTCGGT 2155
Qy 2140 GTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTG 2199
Db 2156 GTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTG 2215
Qy 2200 CGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACT 2259
Db 2216 CGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACT 2275
Qy 2260 GGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTT 2319
Db 2276 GGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTT 2335
Qy 2320 CTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCT 2379

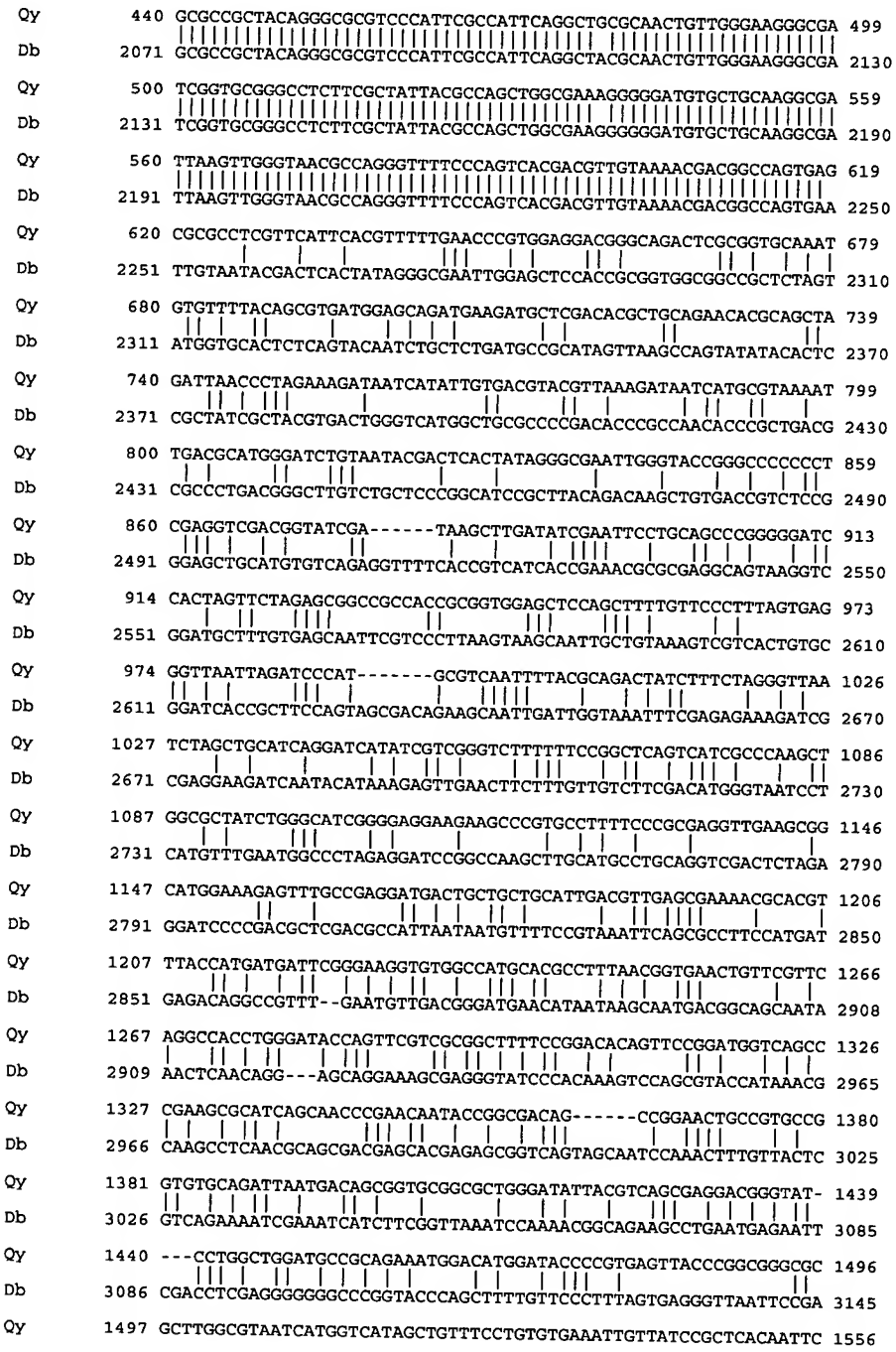
```
Db       2336 CTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCT 2395

Qy       2380 GCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCAC 2439
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2396 GCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCAC 2455

Qy       2440 CGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATC 2499
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2456 CGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATC 2515

Qy       2500 TCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACG 2559
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2516 TCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACG 2575

Qy       2560 TTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTA 2619
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2576 TTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTA 2635

Qy       2620 AAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCA 2679
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2636 AAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCA 2695

Qy       2680 ATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGC 2739
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2696 ATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGC 2755

Qy       2740 CTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGC 2799
             |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2756 CTGACTGCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGC 2815

Qy       2800 TGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCC 2859
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2816 TGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCC 2875

Qy       2860 AGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTAT 2919
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2876 AGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTAT 2935

Qy       2920 TAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGT 2979
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2936 TAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGT 2995

Qy       2980 TGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTC 3039
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2996 TGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTC 3055

Qy       3040 CGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAG 3099
             |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db       3056 CGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGAAAAAAAGCGGTTAG 3115

Qy       3100 CTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGT 3159
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db       3116 CTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGCT 3175

Qy       3160 TATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGAC 3219
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3176 TATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGAC 3235

Qy       3220 TGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTG 3279
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3236 TGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTG 3295

Qy       3280 CCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCAT 3339
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3296 CCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCAT 3355

Qy       3340 TGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTC 3399
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3356 TGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTC 3415

Qy       3400 GATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTC 3459
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db       3416 GATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTC 3475

Qy       3460 TGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAA 3519
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3476 TGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAA 3535

Qy       3520 ATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTG 3579
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3536 ATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTG 3595

Qy       3580 TCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCG 3639
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3596 TCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCG 3655

Qy       3640 CACATTTCCCCGAAAAGTGCCAC 3662
             |||||||||||||||||||||||
Db       3656 CACATTTCCCCGAAAAGTGCCAC 3678
```

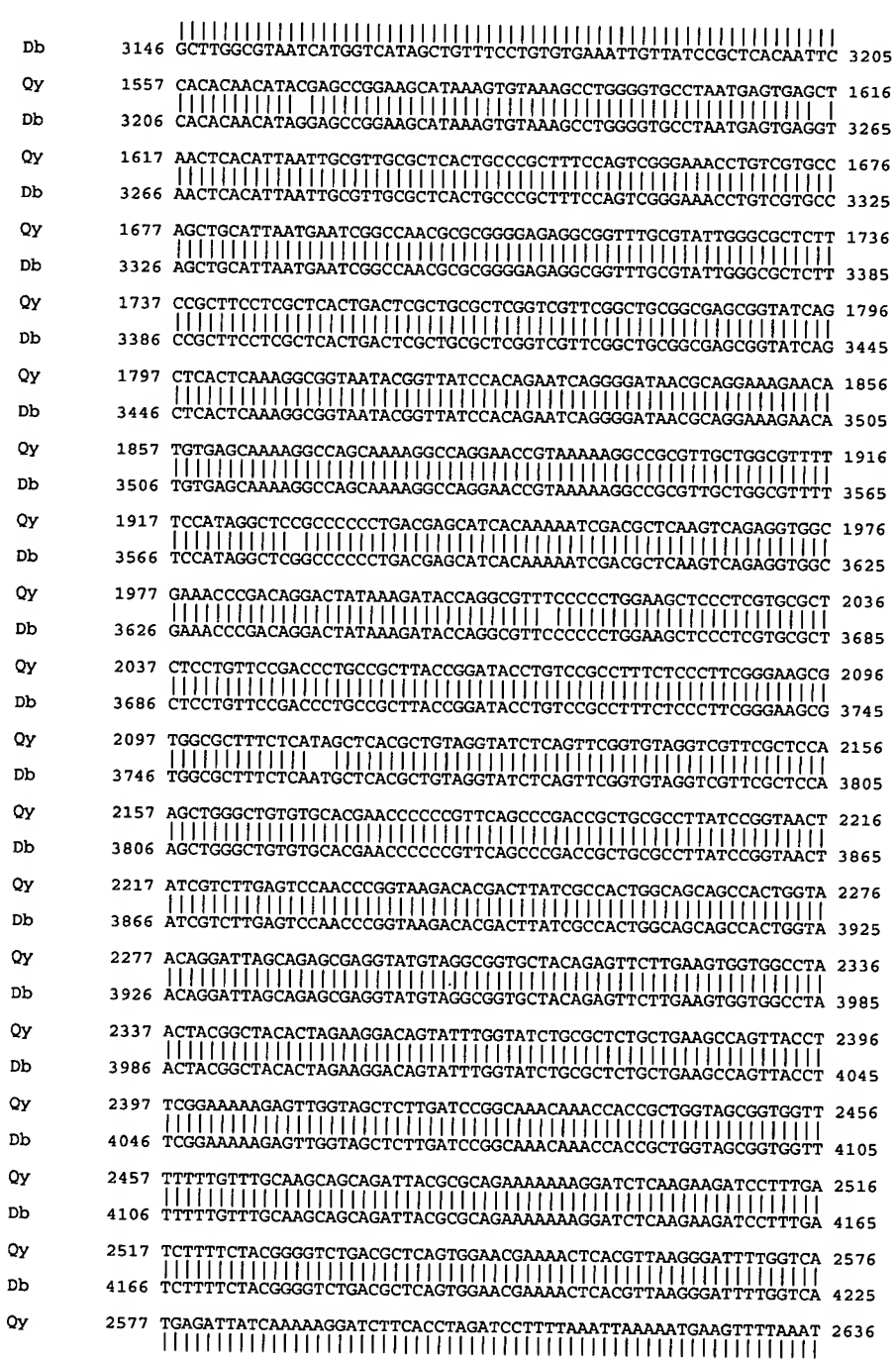
RESULT 7
A20700

```
LOCUS       A20700                  5314 bp    DNA     linear   PAT 22-AUG-1994
DEFINITION  pMTV1 DNA sequence.
ACCESSION   A20700
VERSION     A20700.1  GI:583402
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 5314)
  AUTHORS   .
  TITLE     CARRIER-BOUND RECOMBINANT PROTEIN, PROCESS FOR PRODUCING IT AND ITS
            USE AS AN IMMUNOGEN AND VACCINE
  JOURNAL   Patent: WO 9113155-A 4 05-SEP-1991;
FEATURES             Location/Qualifiers
     source          1. .5314
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
ORIGIN
```

Query Match 71.6%; Score 2623.6; DB 6; Length 5314;
Best Local Similarity 83.4%; Pred. No. 0;
Matches 3059; Conservative 0; Mismatches 579; Indels 28; Gaps 6;

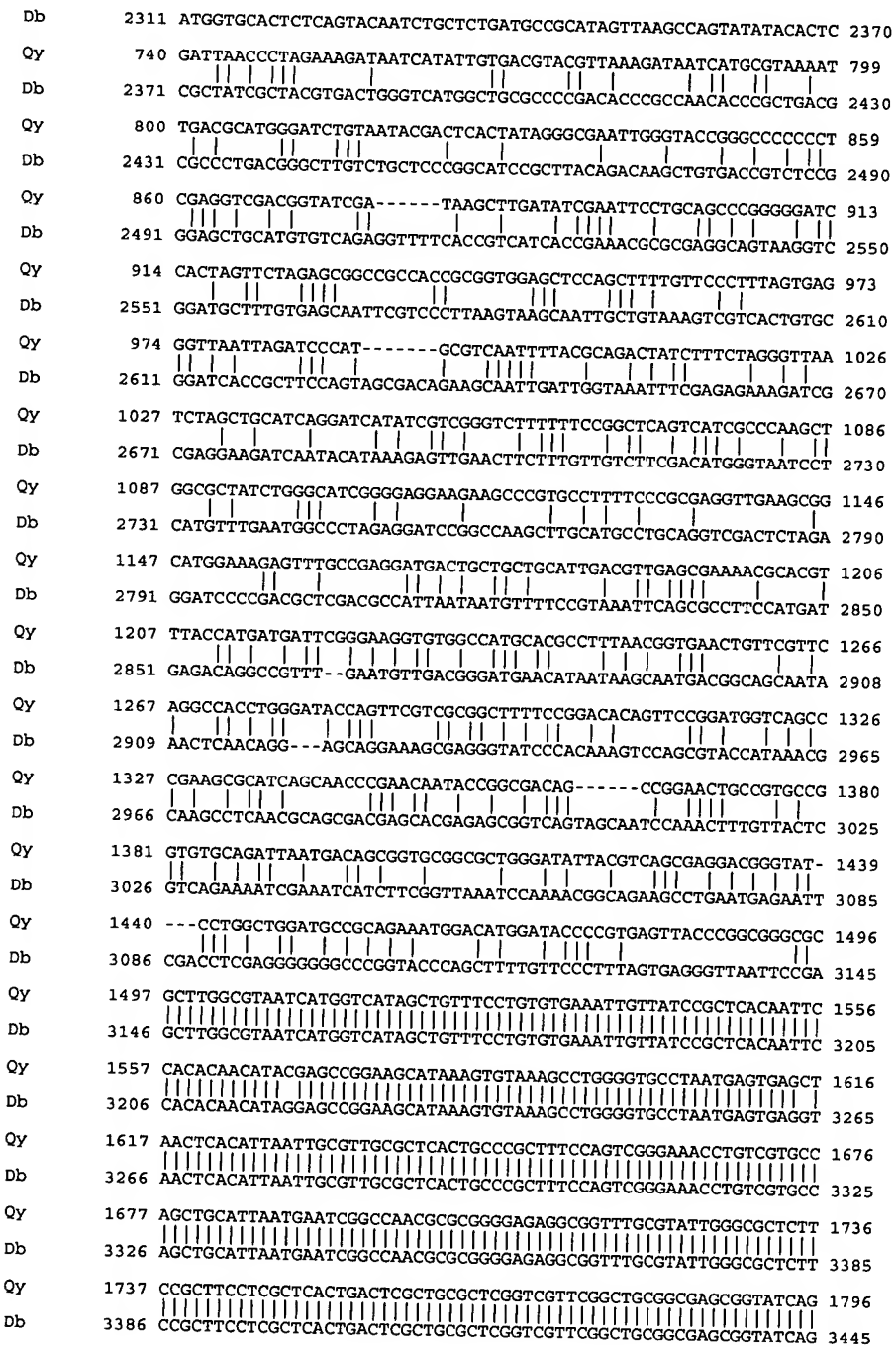
```
Qy         20 ATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTCATTTTTTAACCAATAGGCC 79
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1651 ATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTCATTTTTTAACCAATAGGCC 1710

Qy         80 GAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGAGATAGGGTTGAGTGTTGTT 139
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1711 GAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGAGATAGGGTTGAGTGTTGTT 1770

Qy        140 CCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTCCAACGTCAAAGGGCGAAAA 199
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1771 CCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTCCAACGTCAAAGGGCGAAAA 1830

Qy        200 ACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCCTAATCAAGTTTTTTGGGG 259
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1831 ACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCCTAATCAAGTTTTTTGGGG 1890

Qy        260 TCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAGCCCCCGATTTAGAGCTTGA 319
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1891 TCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAGCCCCCGATTTAGAGCTTGA 1950

Qy        320 CGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAAAGCGAAAGGAGCGGGCGCT 379
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1951 CGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAAAGCGAAAGGAGCGGGCGCT 2010

Qy        380 AGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCACCACACCCGCCGCGCTTAAT 439
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2011 AGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCACCACACCCGCCGCGCTTAAT 2070
```





```
Db       4226 TGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAAT 4285

Qy       2637 CAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGG 2696
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4286 CAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGG 4345

Qy       2697 CACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGT 2756
              |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db       4346 CACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTGCCCGTCGTGT 4405

Qy       2757 AGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAG 2816
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4406 AGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAG 4465

Qy       2817 ACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGC 2876
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4466 ACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGC 4525

Qy       2877 GCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAG 2936
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4526 GCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAG 4585

Qy       2937 CTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCA 2996
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4586 CTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCA 4645

Qy       2997 TCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAA 3056
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4646 TCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAA 4705

Qy       3057 GGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGA 3116
              |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db       4706 GGCGAGTTACATGATCCCCCATGTTGTGAAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGA 4765

Qy       3117 TCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATA 3176
              |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db       4766 TCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGCTTATGGCAGCACTGCATA 4825

Qy       3177 ATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCA 3236
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4826 ATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCA 4885

Qy       3237 AGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGG 3296
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4886 AGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGG 4945

Qy       3297 ATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGG 3356
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4946 ATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGG 5005

Qy       3357 GGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTG 3416
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5006 GGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTG 5065

Qy       3417 CACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAG 3476
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5066 CACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAG 5125

Qy       3477 GAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATAC 3536
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5126 GAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATAC 5185

Qy       3537 TCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACA 3596
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5186 TCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACA 5245

Qy       3597 TATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAG 3656
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5246 TATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAG 5305

Qy       3657 TGCCAC 3662
              ||||||
Db       5306 TGCCAC 5311
```

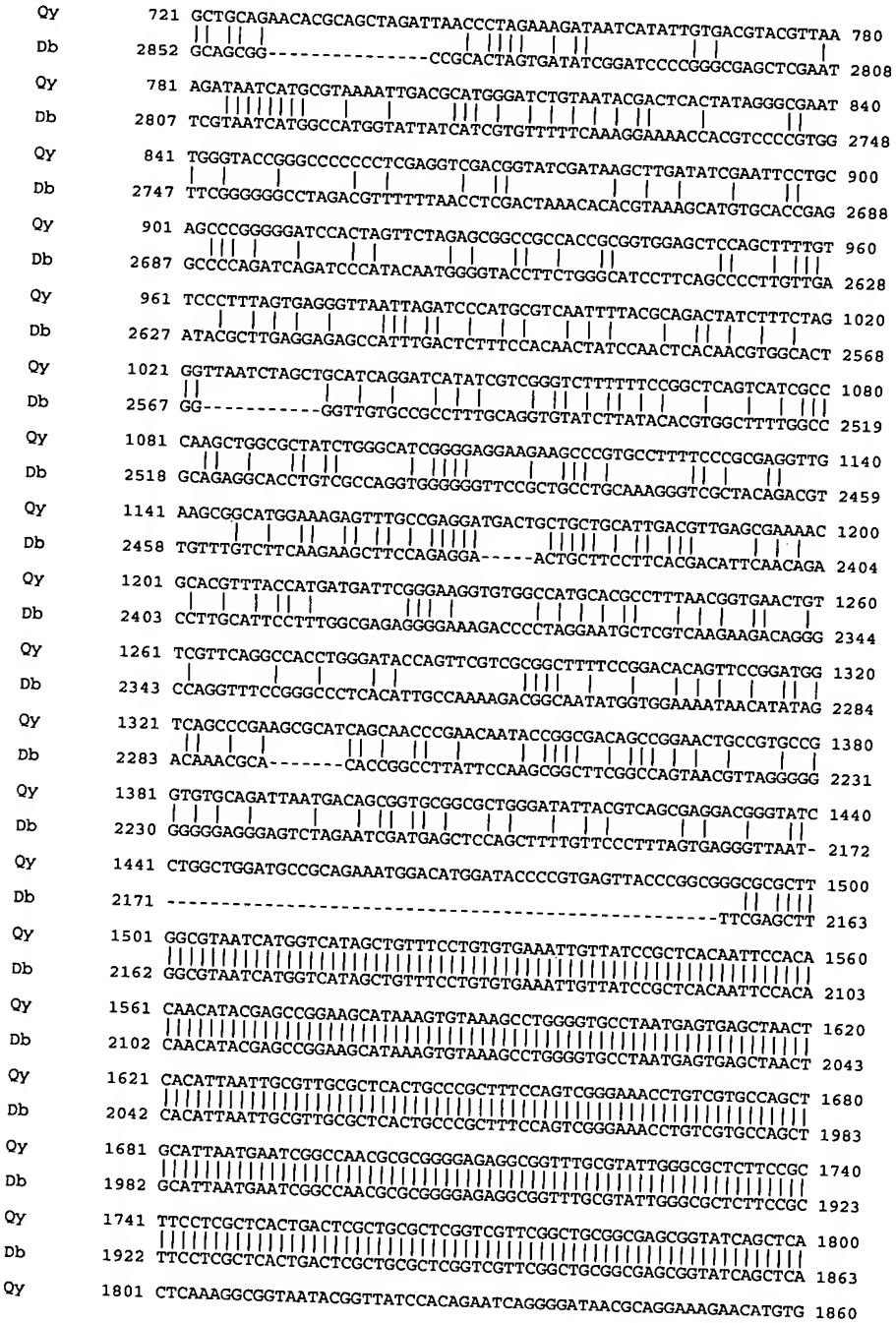
```
RESULT 8
I15649
LOCUS       I15649                5314 bp    DNA     linear   PAT 02-APR-1996
DEFINITION  Sequence 4 from patent US 5470573.
ACCESSION   I15649
VERSION     I15649.1  GI:1250557
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 5314)
  AUTHORS   Lubitz,W. and Szostak,M.P.
  TITLE     Immunogens comprising the non-lytic membrane spanning domain of
            bacteriophages MS2 or PhiX174
  JOURNAL   Patent: US 5470573-A 4 28-NOV-1995;
FEATURES             Location/Qualifiers
     source          1. .5314
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match           71.6%;  Score 2623.6;  DB 6;  Length 5314;
  Best Local Similarity   83.4%;  Pred. No. 0;
  Matches 3059;  Conservative    0;  Mismatches  579;  Indels   28;  Gaps    6;

Qy         20 ATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTCATTTTTTAACCAATAGGCC 79
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1651 ATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTCATTTTTTAACCAATAGGCC 1710

Qy         80 GAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGAGATAGGGTTGAGTGTTGTT 139
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1711 GAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGAGATAGGGTTGAGTGTTGTT 1770

Qy        140 CCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTCCAACGTCAAAGGGCGAAAA 199
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1771 CCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTCCAACGTCAAAGGGCGAAAA 1830

Qy        200 ACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCCTAATCAAGTTTTTTGGGG 259
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1831 ACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCCTAATCAAGTTTTTTGGGG 1890

Qy        260 TCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAGCCCCCGATTTAGAGCTTGA 319
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1891 TCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAGCCCCCGATTTAGAGCTTGA 1950

Qy        320 CGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAAAGCGAAAGGAGCGGGCGCT 379
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1951 CGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAAAGCGAAAGGAGCGGGCGCT 2010

Qy        380 AGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCACCACACCCGCCGCGCTTAAT 439
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2011 AGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCACCACACCCGCCGCGCTTAAT 2070

Qy        440 GCGCCGCTACAGGGCGCGTCCCATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGGGCGA 499
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db       2071 GCGCCGCTACAGGGCGCGTCCCATTCGCCATTCAGGCTACGCAACTGTTGGGAAGGGCGA 2130

Qy        500 TCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGGGGATGTGCTGCAAGGCGA 559
              |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db       2131 TCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAGGGGGGATGTGCTGCAAGGCGA 2190

Qy        560 TTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTGTAAAACGACGGCCAGTGAG 619
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2191 TTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTGTAAAACGACGGCCAGTGAA 2250

Qy        620 CGCGCCTCGTTCATTCACGTTTTTGAACCCGTGGAGGACGGGCAGACTCGCGGTGCAAAT 679
                  |   |   |     |||  | |  ||| |       || |  | | |
Db       2251 TTGTAATACGACTCACTATAGGGCGAATTGGAGCTCCACCGCGGTGGCGGCCGCTCTAGT 2310

Qy        680 GTGTTTTACAGCGTGATGGAGCAGATGAAGATGCTCGACACGCTGCAGAACACGCAGCTA 739
               || |  ||   .  |   | | | |     | |      ||          ||
```

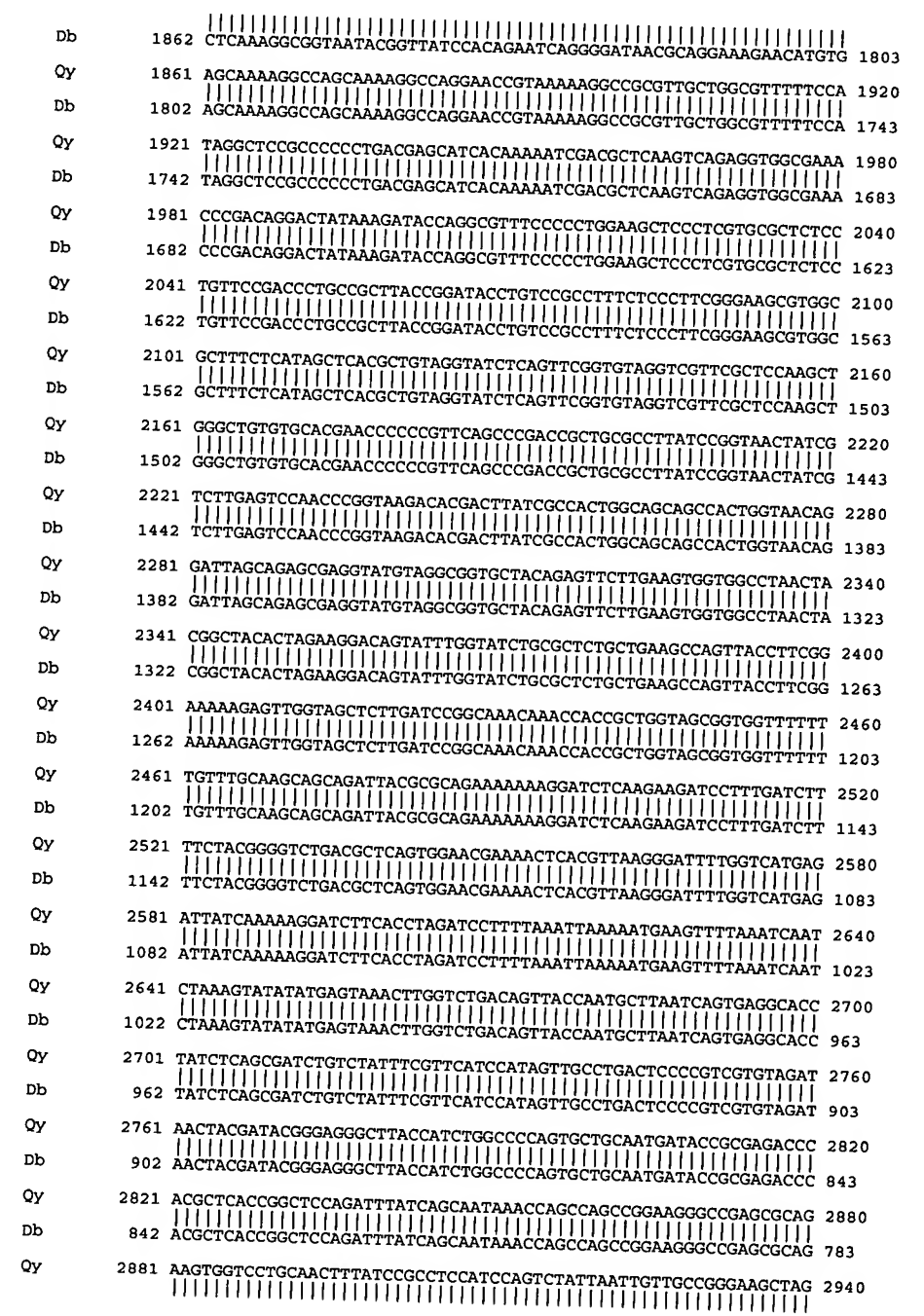

Db 2311 ATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGTATATACACTC 2370
Qy 740 GATTAACCCTAGAAAGATAATCATATTGTGACGTACGTTAAAGATAATCATGCGTAAAAT 799
Db 2371 CGCTATCGCTACGTGACTGGGTCATGGCTGCGCCCCGACACCCGCCAACACCCGCTGACG 2430
Qy 800 TGACGCATGGGATCTGTAATACGACTCACTATAGGGCGAATTGGGTACCGGGCCCCCCCCT 859
Db 2431 CGCCCTGACGGGCTTGTCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCG 2490
Qy 860 CGAGGTCGACGGTATCGA------TAAGCTTGATATCGAATTCCTGCAGCCCGGGGGATC 913
Db 2491 GGAGCTGCATGTGTCAGAGGTTTTCACCGTCATCACCGAAACGCGCGAGGCAGTAAGGTC 2550
Qy 914 CACTAGTTCTAGAGCGGCCGCCACCGCGGTGGAGCTCCAGCTTTTGTTCCCTTTAGTGAG 973
Db 2551 GGATGCTTTGTGAGCAATTCGTCCCTTAAGTAAGCAATTGCTGTAAAGTCGTCACTGTGC 2610
Qy 974 GGTTAATTAGATCCCAT-------GCGTCAATTTTACGCAGACTATCTTTCTAGGGTTAA 1026
Db 2611 GGATCACCGCTTCCAGTAGCGACAGAAGCAATTGATTGGTAAATTTCGAGAGAAAGATCG 2670
Qy 1027 TCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGCT 1086
Db 2671 CGAGGAAGATCAATACATAAAGAGTTGAACTTCTTTGTTGTCTTCGACATGGGTAATCCT 2730
Qy 1087 GGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCGG 1146
Db 2731 CATGTTTGAATGGCCCTAGAGGATCCGGCCAAGCTTGCATGCCTGCAGGTCGACTCTAGA 2790
Qy 1147 CATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACGT 1206
Db 2791 GGATCCCCGACGCTCGACGCCATTAATAATGTTTTCCGTAAATTCAGCGCCTTCCATGAT 2850
Qy 1207 TTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTTC 1266
Db 2851 GAGACAGGCCGTTT--GAATGTTGACGGGATGAACATAATAAGCAATGACGGCAGCAATA 2908
Qy 1267 AGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGCC 1326
Db 2909 AACTCAACAGG---AGCAGGAAAGCGAGGGTATCCCACAAAGTCCAGCGTACCATAAACG 2965
Qy 1327 CGAAGCGCATCAGCAACCCGAACAATACCGGCGACAG------CCGGAACTGCCGTGCCG 1380
Db 2966 CAAGCCTCAACGCAGCGACGAGCACGAGAGCGGTCAGTAGCAATCCAAACTTTGTTACTC 3025
Qy 1381 GTGTGCAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTAT- 1439
Db 3026 GTCAGAAAATCGAAATCATCTTCGGTTAAATCCAAAACGGCAGAAGCCTGAATGAGAATT 3085
Qy 1440 ---CCTGGCTGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGC 1496
Db 3086 CGACCTCGAGGGGGGGCCCGGTACCCAGCTTTTGTTCCCTTTAGTGAGGGTTAATTCCGA 3145
Qy 1497 GCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTC 1556
Db 3146 GCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTC 3205
Qy 1557 CACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCT 1616
Db 3206 CACACAACATAGGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGGT 3265
Qy 1617 AACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCC 1676
Db 3266 AACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCC 3325
Qy 1677 AGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTT 1736
Db 3326 AGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTT 3385
Qy 1737 CCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAG 1796
Db 3386 CCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAG 3445
Qy 1797 CTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACA 1856
Db 3446 CTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACA 3505
Qy 1857 TGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTT 1916
Db 3506 TGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTT 3565
Qy 1917 TCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGC 1976
Db 3566 TCCATAGGCTCGGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGC 3625
Qy 1977 GAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCT 2036
Db 3626 GAAACCCGACAGGACTATAAAGATACCAGGCGTTCCCCCCTGGAAGCTCCCTCGTGCGCT 3685
Qy 2037 CTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCG 2096
Db 3686 CTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCG 3745
Qy 2097 TGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCA 2156
Db 3746 TGGCGCTTTCTCAATGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCA 3805
Qy 2157 AGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACT 2216
Db 3806 AGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACT 3865
Qy 2217 ATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTA 2276
Db 3866 ATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTA 3925
Qy 2277 ACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTA 2336
Db 3926 ACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTA 3985
Qy 2337 ACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCT 2396
Db 3986 ACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCT 4045
Qy 2397 TCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTT 2456
Db 4046 TCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTT 4105
Qy 2457 TTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGA 2516
Db 4106 TTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGA 4165
Qy 2517 TCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCA 2576
Db 4166 TCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCA 4225
Qy 2577 TGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAAT 2636
Db 4226 TGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAAT 4285
Qy 2637 CAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGG 2696
Db 4286 CAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGG 4345
Qy 2697 CACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGT 2756
Db 4346 CACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTGCCCGTCGTGT 4405
Qy 2757 AGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAG 2816
Db 4406 AGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAG 4465
Qy 2817 ACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGC 2876
Db 4466 ACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGC 4525

```
Qy       2877 GCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAG 2936
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4526 GCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAG 4585

Qy       2937 CTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCA 2996
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4586 CTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCA 4645

Qy       2997 TCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAA 3056
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4646 TCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAA 4705

Qy       3057 GGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGA 3116
             |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db       4706 GGCGAGTTACATGATCCCCCATGTTGTGAAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGA 4765

Qy       3117 TCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATA 3176
             |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db       4766 TCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGCTTATGGCAGCACTGCATA 4825

Qy       3177 ATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCA 3236
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4826 ATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCA 4885

Qy       3237 AGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGG 3296
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4886 AGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGG 4945

Qy       3297 ATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGG 3356
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4946 ATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGG 5005

Qy       3357 GGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTG 3416
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5006 GGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTG 5065

Qy       3417 CACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAG 3476
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5066 CACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAG 5125

Qy       3477 GAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATAC 3536
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5126 GAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATAC 5185

Qy       3537 TCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACA 3596
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5186 TCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACA 5245

Qy       3597 TATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAG 3656
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5246 TATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAG 5305

Qy       3657 TGCCAC 3662
             ||||||
Db       5306 TGCCAC 5311
```

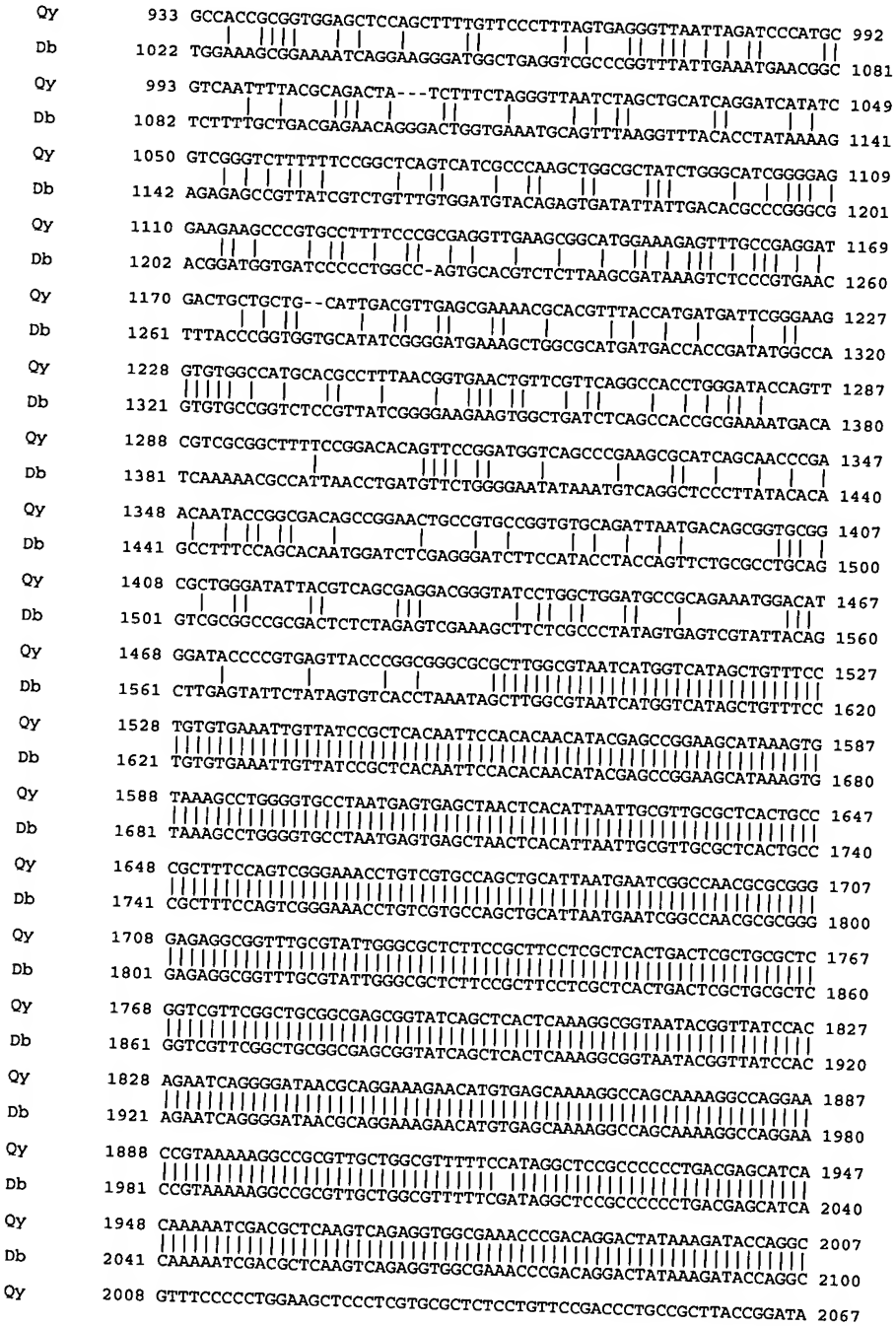
```
RESULT 9
AF171227/c
LOCUS       AF171227               3552 bp    DNA     circular SYN 04-FEB-2000
DEFINITION  Shuttle vector pSLIRES11, complete sequence.
ACCESSION   AF171227
VERSION     AF171227.1  GI:5764544
KEYWORDS    .
SOURCE      Shuttle vector pSLIRES11
  ORGANISM  Shuttle vector pSLIRES11
            artificial sequences; vectors.
REFERENCE   1  (bases 1 to 3552)
  AUTHORS   Chen,C.M., Smith,D.M., Peters,M.A., Samson,M.E., Zitz,J.,
            Tabin,C.J. and Cepko,C.L.
  TITLE     Production and design of more effective avian
            replication-incompetent retroviral vectors
  JOURNAL   Dev. Biol. 214 (2), 370-384 (1999)
  MEDLINE   99456779
   PUBMED   10525341
REFERENCE   2  (bases 1 to 3552)
  AUTHORS   Chen,C.-M.A., Samson,M.E.S. and Cepko,C.L.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JUL-1999) Genetics, Harvard Medical School, 200
            Longwood Avenue, Boston, MA 02115, USA
FEATURES             Location/Qualifiers
     source          1. .3552
                     /organism="Shuttle vector pSLIRES11"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:101936"
                     /lab_host="Escherichia coli"
     misc_feature    join(2864. .3552,1. .2208)
                     /note="similar to pBluescript"
     misc_feature    2236. .2789
                     /note="IRES of EMCV"
     misc_feature    2790. .2849
                     /note="multiple cloning site"
ORIGIN
```

```
  Query Match           71.3%;  Score 2610.8;  DB 12;  Length 3552;
  Best Local Similarity   83.8%;  Pred. No. 0;
  Matches 3070;  Conservative    0;  Mismatches  482;  Indels  110;  Gaps    6;

Qy          1 CTAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3552 CTAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTC 3493

Qy         61 ATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3492 ATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGA 3433

Qy        121 GATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3432 GATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTC 3373

Qy        181 CAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3372 CAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACC 3313

Qy        241 CTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3312 CTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAG 3253

Qy        301 CCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3252 CCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAA 3193

Qy        361 AGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCAC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3192 AGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCAC 3133

Qy        421 CACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCCCATTCGCCATTCAGGCTGCG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3132 CACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCCCATTCGCCATTCAGGCTGCG 3073

Qy        481 CAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3072 CAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGG 3013

Qy        541 GGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3012 GGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTG 2953

Qy        601 TAAAACGACGGCCAGTGAGCGCGCCTCGTTCATTCACGTTTTTGAACCCGTGGAGGACGG 660
              ||||||||||||||||||                  || | | |    | | |
Db       2952 TAAAACGACGGCCAGTGA------------------ATTGTAATACGACTCACTATAG 2913

Qy        661 GCAGACTCGCGGTGCAAATGTGTTTTACAGCGTGATGGAGCAGATGAAGATGCTCGACAC 720
              | || | | |  |           | | || | | ||   ||  | | ||| |
Db       2912 GGCGAATTGGGTACCGGGCCCCCCCTCGATCGAGGTCGACGGTATCGATAAGCTTGGGCT 2853
```

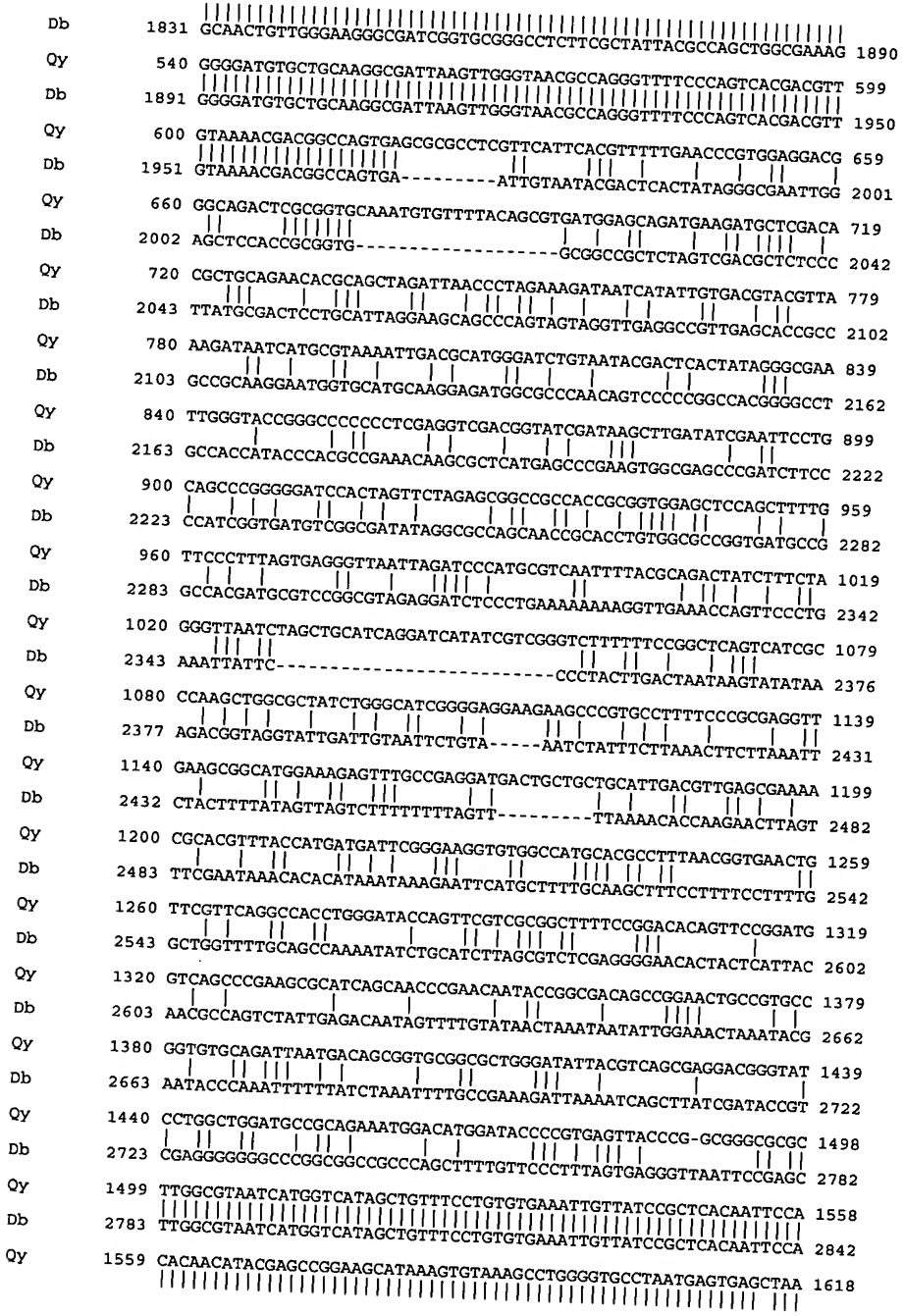




```
Db        782 AAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAG 723

Qy       2941 AGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGT 3000
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        722 AGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGT 663

Qy       3001 GGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCG 3060
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        662 GGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCG 603

Qy       3061 AGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGT 3120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        602 AGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGT 543

Qy       3121 TGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTC 3180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        542 TGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTC 483

Qy       3181 TCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTC 3240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        482 TCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTC 423

Qy       3241 ATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAA 3300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 ATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAA 363

Qy       3301 TACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCG 3360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 TACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCG 303

Qy       3361 AAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACC 3420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 AAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACC 243

Qy       3421 CAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAG 3480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 CAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAG 183

Qy       3481 GCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTT 3540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 GCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTT 123

Qy       3541 CCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATT 3600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 CCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATT 63

Qy       3601 TGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCC 3660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          62 TGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCC 3

Qy       3661 AC 3662
              ||
Db          2 AC 1


RESULT 10
AX306330
LOCUS       AX306330               4001 bp    DNA     linear   PAT 11-DEC-2001
DEFINITION  Sequence 13 from Patent WO0188121.
ACCESSION   AX306330
VERSION     AX306330.1  GI:17645569
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Plaetinck,G., Renard,J.P. and Bogaert,T.
  TITLE     Vector constructs
  JOURNAL   Patent: WO 0188121-A 13 22-NOV-2001;
            Devgen NV (BE)
FEATURES             Location/Qualifiers
     source          1. .4001
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="Plasmid pGN59A"
ORIGIN

  Query Match             71.3%;  Score 2609.4;  DB 6;  Length 4001;
  Best Local Similarity   83.1%;  Pred. No. 0;
  Matches 3069;  Conservative    0;  Mismatches  591;  Indels   35;  Gaps    7;

Qy          1 CTAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTC 60
              | |||||||| ||||||||||||||||||||||||||||||| ||||||||||||||||||
Db         63 CGAAATTGTAAACGTTAATATTTTGTTAAAATTCGCGTTAAATATTTGTTAAATCAGCTC 122

Qy         61 ATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGA 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 ATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGA 182

Qy        121 GATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTC 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 GATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTC 242

Qy        181 CAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACC 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 CAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACC 302

Qy        241 CTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAG 300
              | |||||||||||| |||||||||||||||||||||||| ||||||||||||||||||||
Db        303 CAAATCAAGTTTTTTGCGGTCGAGGTGCCGTAAAGCTCTAAATCGGAACCCTAAAGGGAG 362

Qy        301 CCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAA 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 CCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAA 422

Qy        361 AGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCAC 420
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        423 AGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCAC 482

Qy        421 CACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCCCATTCGCCATTCAGGCTGCG 480
              |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        483 CACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGT-CCATTCGCCATTCAGGCTGCG 541

Qy        481 CAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGG 540
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        542 CAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGG 601

Qy        541 GGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTG 600
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        602 GGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTG 661

Qy        601 TAAAACGACGGCCAGTGAGCGCGCCT--------------CGTTCATTCACGTTTTTGAA 646
              |||||||||||||||||| |                   |  ||||  | |    |
Db        662 TAAAACGACGGCCAGTGAATTGTAATACGACTCACTATAGGGCGAATTCGAGCTCGGTAC 721

Qy        647 CCCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGCGTGATGGAGCAGATG 706
              || | |     | |    | | ||  .  |  |   ||     |  | ||| | |
Db        722 CCGGGGATCCTCTAGAGATCCCTCGACCTCGAGATCCATTGTGCTGGAAAGGATCTGGAT 781

Qy        707 AAGATGCTCGACACGCTGCAGAACACGCAGCTAGAT-------TAACCCTAGAAAGATAA 759
                |  | | | ||  | ||||   ||   |  | |  |  |    ||||
Db        782 CCGGCTTACTAAAAGCCAGATAACAGTATGCGTATTTGCGCGCTGATTTTTGCGGTATAA 841

Qy        760 TCATATTGTGACGTACGTTAAAGATAATCATGCGTAAAATTG-------ACGCATGGGAT 812
              ||||        || ||  |   ||  |||   ||||  |       | | |   |
Db        842 GAATATATACTGATATGTATACCCGAAGTATGTCAAAAAGAGGTGTGCTATGAAGCAGCG 901

Qy        813 CTGTAATACGACTCACTATAGGGCGAATTGGGTACCGGGCCCCCCCTCGAGGTCGACGGT 872
              ||   |||     | || | |  |      |  |   | |  | | | | | |  |
Db        902 TATTACAGTGACAGTTGACAGCGACAGCTATCAGTTGCTCAAGGCATATATGATGTCAAT 961

Qy        873 ATCGATAAGCTTGATATCGAATTCCTGCAGCCCGGGGGATCCACTAGTTCTAGAGCGGCC 932
              |||   || | | | |  | |||||  |  |     |  |  | | |  |
Db        962 ATCTCCGGTCTGGTAAGCACAACCATGCAGAATGAAGCCCGTCGTCTGCGTGCCGAACGC 1021
```


Qy 933 GCCACCGCGGTGGAGCTCCAGCTTTTGTTCCCTTTAGTGAGGGTTAATTAGATCCCATGC 992
Db 1022 TGGAAAGCGGAAAATCAGGAAGGGATGGCTGAGGTCGCCCGGTTTATTGAAATGAACGGC 1081
Qy 993 GTCAATTTTACGCAGACTA---TCTTTCTAGGGTTAATCTAGCTGCATCAGGATCATATC 1049
Db 1082 TCTTTTGCTGACGAGAACAGGGACTGGTGAAATGCAGTTTAAGGTTTACACCTATAAAAG 1141
Qy 1050 GTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGGCATCGGGGAG 1109
Db 1142 AGAGAGCCGTTATCGTCTGTTTGTGGATGTACAGAGTGATATTATTGACACGCCCGGGCG 1201
Qy 1110 GAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCGGCATGGAAAGAGTTTGCCGAGGAT 1169
Db 1202 ACGGATGGTGATCCCCCTGGCC-AGTGCACGTCTCTTAAGCGATAAAGTCTCCCGTGAAC 1260
Qy 1170 GACTGCTGCTG--CATTGACGTTGAGCGAAAACGCACGTTTACCATGATGATTCGGGAAG 1227
Db 1261 TTTACCCGGTGGTGCATATCGGGGATGAAAGCTGGCGCATGATGACCACCGATATGGCCA 1320
Qy 1228 GTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTTCAGGCCACCTGGGATACCAGTT 1287
Db 1321 GTGTGCCGGTCTCCGTTATCGGGGAAGAAGTGGCTGATCTCAGCCACCGCGAAAATGACA 1380
Qy 1288 CGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGCCCGAAGCGCATCAGCAACCCGA 1347
Db 1381 TCAAAAACGCCATTAACCTGATGTTCTGGGGAATATAAATGTCAGGCTCCCTTATACACA 1440
Qy 1348 ACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTGCAGATTAATGACAGCGGTGCGG 1407
Db 1441 GCCTTTCCAGCACAATGGATCTCGAGGGATCTTCCATACCTACCAGTTCTGCGCCTGCAG 1500
Qy 1408 CGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGCTGGATGCCGCAGAAATGGACAT 1467
Db 1501 GTCGCGGCCGCGACTCTCTAGAGTCGAAAGCTTCTCGCCCTATAGTGAGTCGTATTACAG 1560
Qy 1468 GGATACCCCGTGAGTTACCCGGCGGGCGCGCTTGGCGTAATCATGGTCATAGCTGTTTCC 1527
Db 1561 CTTGAGTATTCTATAGTGTCACCTAAATAGCTTGGCGTAATCATGGTCATAGCTGTTTCC 1620
Qy 1528 TGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTG 1587
Db 1621 TGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTG 1680
Qy 1588 TAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCC 1647
Db 1681 TAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCC 1740
Qy 1648 CGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGG 1707
Db 1741 CGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGG 1800
Qy 1708 GAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTC 1767
Db 1801 GAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTC 1860
Qy 1768 GGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCAC 1827
Db 1861 GGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCAC 1920
Qy 1828 AGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAA 1887
Db 1921 AGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAA 1980
Qy 1888 CCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCA 1947
Db 1981 CCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCGATAGGCTCCGCCCCCCTGACGAGCATCA 2040
Qy 1948 CAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGC 2007
Db 2041 CAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGC 2100
Qy 2008 GTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATA 2067
Db 2101 GTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATA 2160
Qy 2068 CCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTA 2127
Db 2161 CCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTA 2220
Qy 2128 TCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCA 2187
Db 2221 TCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCA 2280
Qy 2188 GCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGA 2247
Db 2281 GCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGA 2340
Qy 2248 CTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGG 2307
Db 2341 CTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGG 2400
Qy 2308 TGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGG 2367
Db 2401 TGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGG 2460
Qy 2368 TATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGG 2427
Db 2461 TATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGG 2520
Qy 2428 CAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAG 2487
Db 2521 CAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAG 2580
Qy 2488 AAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAA 2547
Db 2581 AAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAA 2640
Qy 2548 CGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGAT 2607
Db 2641 CGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGAT 2700
Qy 2608 CCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTC 2667
Db 2701 CCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTC 2760
Qy 2668 TGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTC 2727
Db 2761 TGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTC 2820
Qy 2728 ATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATC 2787
Db 2821 ATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATC 2880
Qy 2788 TGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGC 2847
Db 2881 TGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGC 2940
Qy 2848 AATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTC 2907
Db 2941 AATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTC 3000
Qy 2908 CATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTT 2967
Db 3001 CATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTT 3060
Qy 2968 GCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGC 3027
Db 3061 GCGCAACGTTGTTGGCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGC 3120
Qy 3028 TTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAA 3087
Db 3121 TTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAA 3180
Qy 3088 AAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTT 3147

```
Db       3181 AAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTT 3240

Qy       3148 ATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATG 3207
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3241 ATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATG 3300

Qy       3208 CTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACC 3267
              ||||||||||||||||||||||||||||||||||||||||||  |    ||||||||
Db       3301 CTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATACCGCGCCCGGCGACC 3360

Qy       3268 GAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAA 3327
              ||||||||||||||||||||||||||||||||  |    ||||||||||||||||||
Db       3361 GAGTTGCTCTTGCCCGGCGTCAATACGGGATAATAGTGTATGACATAGCAGAACTTTAAA 3420

Qy       3328 AGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTT 3387
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3421 AGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTT 3480

Qy       3388 GAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTT 3447
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3481 GAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTT 3540

Qy       3448 CACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAG 3507
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3541 CACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAG 3600

Qy       3508 GGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTA 3567
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3601 GGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTA 3660

Qy       3568 TCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAAT 3627
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3661 TCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAAT 3720

Qy       3628 AGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCAC 3662
              |||||||||||||||||||||||||||||||||||
Db       3721 AGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCAC 3755


RESULT 11
AF179627
LOCUS       AF179627               5540 bp    DNA     linear   SYN 22-NOV-1999
DEFINITION  Expression vector pCENEX645, complete sequence.
ACCESSION   AF179627
VERSION     AF179627.1  GI:6457303
KEYWORDS    .
SOURCE      Expression vector pCENEX645
  ORGANISM  Expression vector pCENEX645
            artificial sequences; vectors.
REFERENCE   1  (bases 1 to 5540)
  AUTHORS   Eugster,H.P., Bartsch,S., Wurgler,F.E. and Sengstag,C.
  TITLE     Functional co-expression of human oxidoreductase and cytochrome
            P450 1A1 in Saccharomyces cerevisiae results in increased EROD
            activity
  JOURNAL   Biochem. Biophys. Res. Commun. 185 (2), 641-647 (1992)
  MEDLINE   92304288
   PUBMED   1610357
REFERENCE   2  (bases 1 to 5540)
  AUTHORS   Sengstag,C. and Paladino,G.
  TITLE     A collection of transformed Saccharomyces cerevisiae strains as an
            unlimited source of human enzymes
  JOURNAL   BioWorld 97 (6), 2-7 (1997)
REFERENCE   3  (bases 1 to 5540)
  AUTHORS   Sengstag,C.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-AUG-1999) Genetics, Institute of Toxicology, ETH
            Zurich, Schorenstrasse 16, Schwerzenbach 8603, Switzerland
FEATURES             Location/Qualifiers
     source          1. .5540
                     /organism="Expression vector pCENEX645"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:106613"
                     /note="yeast expression vector pCENEX645 (alias pCS645)
                     containing truncated GAPDH promoter for constitutive
                     expression of cDNAs in yeast"
     gene            417. .1220
                     /gene="URA3"
     CDS             417. .1220
                     /gene="URA3"
                     /note="yeast selection marker"
                     /codon_start=1
                     /product="URA3"
                     /protein_id="AAF09466.1"
                     /db_xref="GI:6457304"
                     /translation="MSKATYKERAATHPSPVAAKLFNIMHEKQTNLCASLDVRTTKEL
                     LELVEALGPKICLLKTHVDILTDFSMEGTVKPLKALSAKYNFLLFEDRKFADIGNTVK
                     LQYSAGVYRIAEWADITNAHGVVGPGIVSGLKQAAEEVTKEPRGLLMLAELSCKGSLS
                     TGEYTKGTVDIAKSDKDFVIGFIAQRDMGGRDEGYDWLIMTPGVGLDDKGDALGQQYR
                     TVDDVVSTGSDIIIVGRGLFAKGRDAKVEGERYRKAGWEAYLRRCGQQN"
     promoter        2310. .2508
                     /note="truncated GAPDH promoter"
     gene            complement(3958. .4818)
                     /gene="Amp"
     CDS             complement(3958. .4818)
                     /gene="Amp"
                     /note="ampicillin resistance; Escherichia coli selection
                     marker"
                     /codon_start=1
                     /product="Amp"
                     /protein_id="AAF09467.1"
                     /db_xref="GI:6457305"
                     /translation="MSIQHFRVALIPFFAAFCLPVFAHFDLLVFVKDCEDQLGARVGY
                     IELDLNSGKILESFRPEERFPMMSTFKVLLCGAVLFRIDAGQEQLGRRIHYSQNDLVE
                     YSPVTEKHLTDGMTVRELCSAAITMSDNTAANLLLTTIGGPKELTAFLHNMGDHVTRL
                     DRWEPELNEAIPNDERDTTMPVAMATTLRKLLTGELLTLASRQQLIDWMEADKVAGPL
                     LRSALPAGWFIADKSGAGERGSRGIIAALGPDGKPSRIVVIYTTGSQATMDERNRQIA
                     EIGASLIKHW"
ORIGIN

  Query Match           69.9%;  Score 2558.8;  DB 12;  Length 5540;
  Best Local Similarity  82.8%;  Pred. No. 0;
  Matches 3035;  Conservative    0;  Mismatches  557;  Indels   72;  Gaps    7;

Qy          3 AAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTCAT 62
              ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1351 AAATTGTAAACGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTCAT 1410

Qy         63 TTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGAGA 122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1411 TTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGAGA 1470

Qy        123 TAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTCCA 182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1471 TAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTCCA 1530

Qy        183 ACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCCT 242
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1531 ACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCCT 1590

Qy        243 AATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAGCC 302
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1591 AATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAGCC 1650

Qy        303 CCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAAAG 362
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1651 CCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAAAG 1710

Qy        363 CGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCACCA 422
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db       1711 CGAAAGGAGCGGGCGCTAGGCCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCACCA 1770

Qy        423 CACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTC---CCATTCGCCATTCAGGCTGC 479
              |||||||||||||||||||||||||||||||||||||   ||||||||||||||||||||
Db       1771 CACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCGCGCCATTCGCCATTCAGGCTGC 1830

Qy        480 GCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAG 539
```

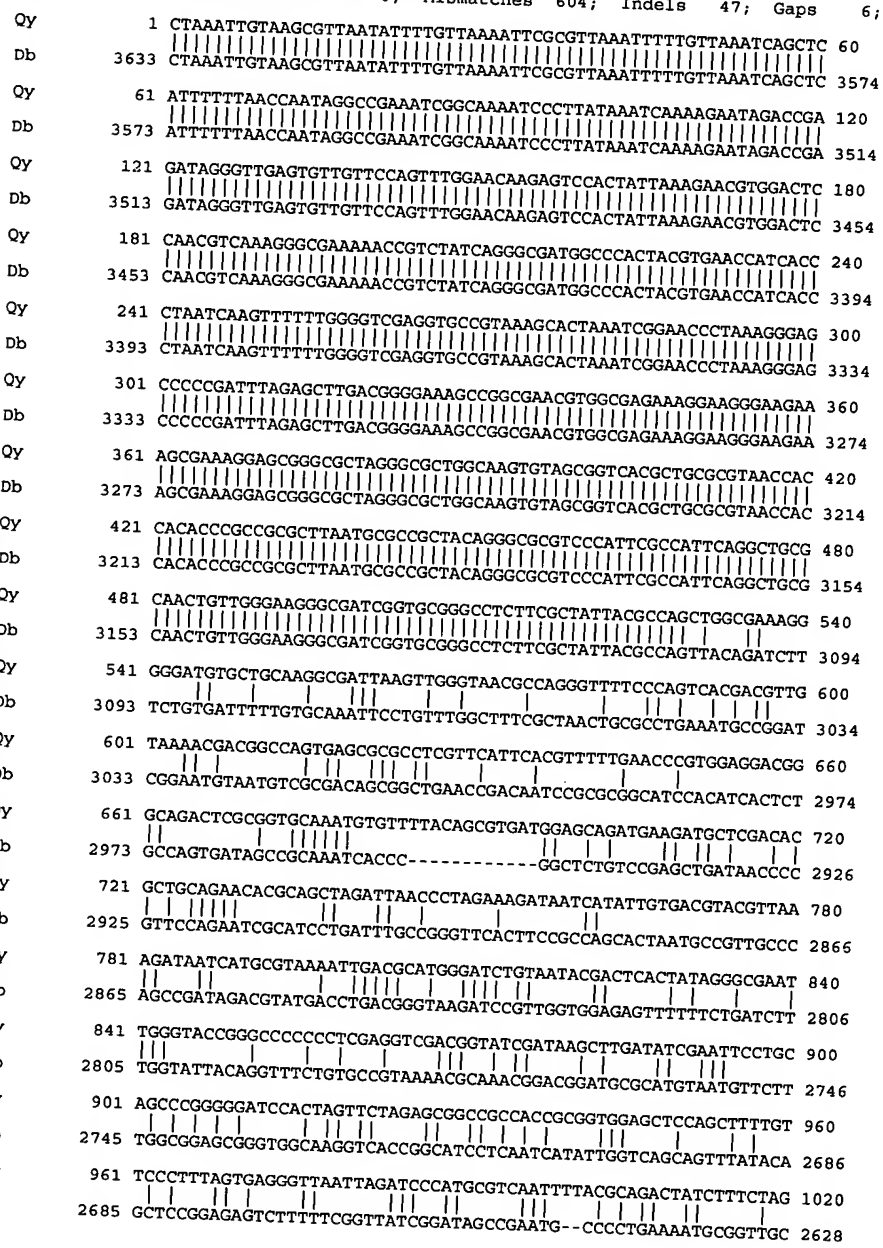

Db 1831 GCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAG 1890
Qy 540 GGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTT 599
Db 1891 GGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTT 1950
Qy 600 GTAAAACGACGGCCAGTGAGCGCGCCTCGTTCATTCACGTTTTTGAACCCGTGGAGGACG 659
Db 1951 GTAAAACGACGGCCAGTGA---------ATTGTAATACGACTCACTATAGGGCGAATTGG 2001
Qy 660 GGCAGACTCGCGGTGCAAATGTGTTTTACAGCGTGATGGAGCAGATGAAGATGCTCGACA 719
Db 2002 AGCTCCACCGCGGTG----------GCGGCCGCTCTAGTCGACGCTCTCCC 2042
Qy 720 CGCTGCAGAACACGCAGCTAGATTAACCCTAGAAAGATAATCATATTGTGACGTACGTTA 779
Db 2043 TTATGCGACTCCTGCATTAGGAAGCAGCCCAGTAGTAGGTTGAGGCCGTTGAGCACCGCC 2102
Qy 780 AAGATAATCATGCGTAAAATTGACGCATGGGATCTGTAATACGACTCACTATAGGGCGAA 839
Db 2103 GCCGCAAGGAATGGTGCATGCAAGGAGATGGCGCCCAACAGTCCCCCGGCCACGGGGCCT 2162
Qy 840 TTGGGTACCGGGCCCCCCCTCGAGGTCGACGGTATCGATAAGCTTGATATCGAATTCCTG 899
Db 2163 GCCACCATACCCACGCCGAAACAAGCGCTCATGAGCCCGAAGTGGCGAGCCCGATCTTCC 2222
Qy 900 CAGCCCGGGGGATCCACTAGTTCTAGAGCGGCCGCCACCGCGGTGGAGCTCCAGCTTTTG 959
Db 2223 CCATCGGTGATGTCGGCGATATAGGCGCCAGCAACCGCACCTGTGGCGCCGGTGATGCCG 2282
Qy 960 TTCCCTTTAGTGAGGGTTAATTAGATCCCATGCGTCAATTTTACGCAGACTATCTTTCTA 1019
Db 2283 GCCACGATGCGTCCGGCGTAGAGGATCTCCCTGAAAAAAAAGGTTGAAACCAGTTCCCTG 2342
Qy 1020 GGGTTAATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGC 1079
Db 2343 AAATTATTC----------CCCTACTTGACTAATAAGTATATAA 2376
Qy 1080 CCAAGCTGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTT 1139
Db 2377 AGACGGTAGGTATTGATTGTAATTCTGTA-----AATCTATTTCTTAAACTTCTTAAATT 2431
Qy 1140 GAAGCGGCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAA 1199
Db 2432 CTACTTTTATAGTTAGTCTTTTTTTTAGTT---------TTAAAACACCAAGAACTTAGT 2482
Qy 1200 CGCACGTTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTG 1259
Db 2483 TTCGAATAAACACACATAAATAAAGAATTCATGCTTTTGCAAGCTTTCCTTTTCCTTTTG 2542
Qy 1260 TTCGTTCAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATG 1319
Db 2543 GCTGGTTTTGCAGCCAAAATATCTGCATCTTAGCGTCTCGAGGGGAACACTACTCATTAC 2602
Qy 1320 GTCAGCCCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCC 1379
Db 2603 AACGCCAGTCTATTGAGACAATAGTTTTGTATAACTAAATAATATTGGAAACTAAATACG 2662
Qy 1380 GGTGTGCAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTAT 1439
Db 2663 AATACCCAAATTTTTTATCTAAATTTTGCCGAAAGATTAAAATCAGCTTATCGATACCGT 2722
Qy 1440 CCTGGCTGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCG-GCGGGCGCGC 1498
Db 2723 CGAGGGGGGGCCCGGCGGCCGCCCAGCTTTTGTTCCCTTTAGTGAGGGTTAATTCCGAGC 2782
Qy 1499 TTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCA 1558
Db 2783 TTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCA 2842
Qy 1559 CACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAA 1618
Db 2843 CACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGGTAA 2902
Qy 1619 CTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAG 1678
Db 2903 CTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGRCAG 2962
Qy 1679 CTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCC 1738
Db 2963 CTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCC 3022
Qy 1739 GCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCT 1798
Db 3023 GCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCT 3082
Qy 1799 CACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATG 1858
Db 3083 CACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATG 3142
Qy 1859 TGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTC 1918
Db 3143 TGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTC 3202
Qy 1919 CATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGA 1978
Db 3203 CATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGA 3262
Qy 1979 AACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCT 2038
Db 3263 AACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCT 3322
Qy 2039 CCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTG 2098
Db 3323 CCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTG 3382
Qy 2099 GCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAG 2158
Db 3383 GCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAG 3442
Qy 2159 CTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTAT 2218
Db 3443 CTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTAT 3502
Qy 2219 CGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAAC 2278
Db 3503 CGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAAC 3562
Qy 2279 AGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAAC 2338
Db 3563 AGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAAC 3622
Qy 2339 TACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTC 2398
Db 3623 TACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTC 3682
Qy 2399 GGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTT 2458
Db 3683 GGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTT 3742
Qy 2459 TTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATC 2518
Db 3743 TTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATC 3802
Qy 2519 TTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATG 2578
Db 3803 TTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATG 3862
Qy 2579 AGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCA 2638
Db 3863 AGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCA 3922
Qy 2639 ATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCA 2698
Db 3923 ATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCA 3982

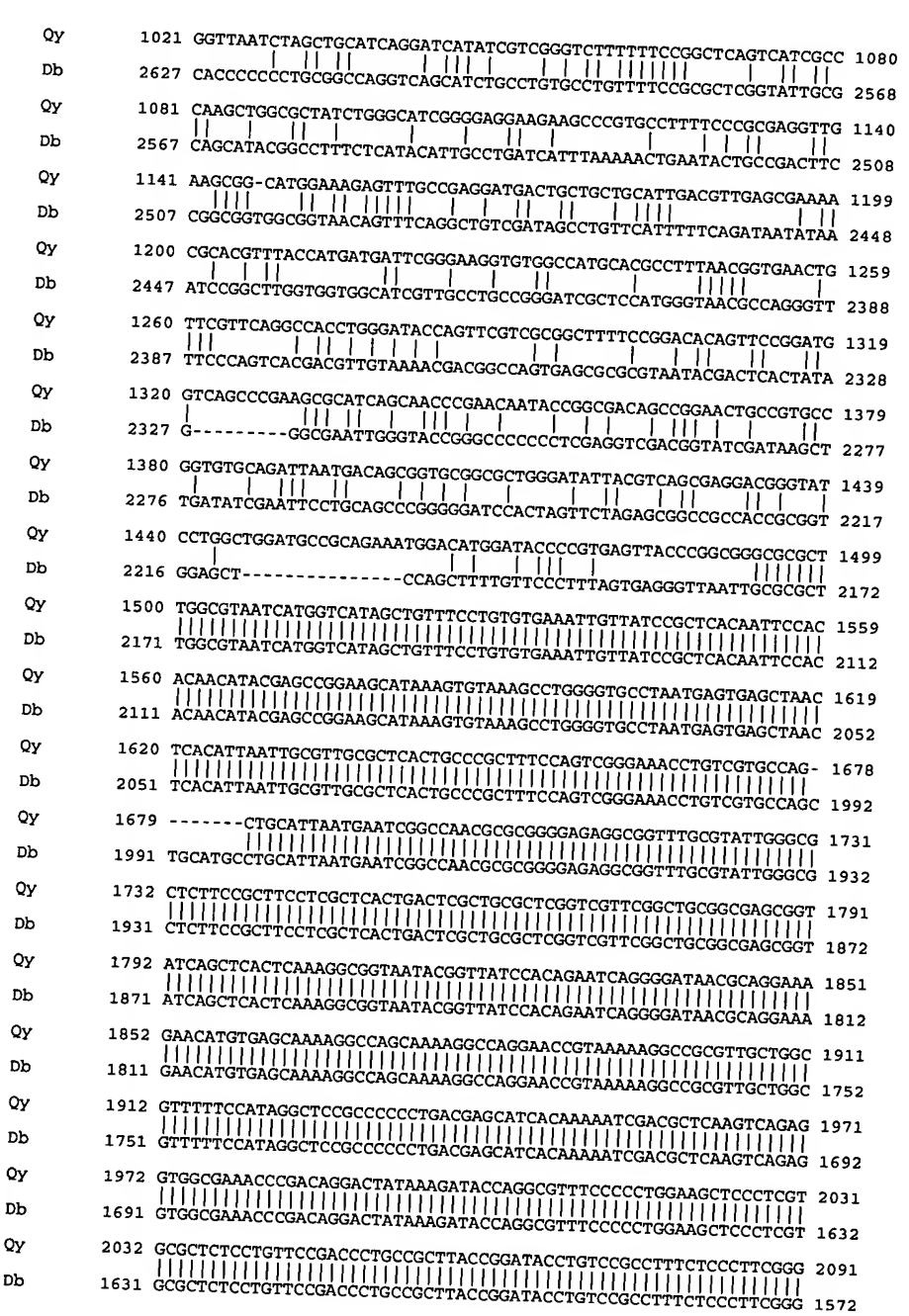
```
Qy       2699 CCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAG 2758
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3983 CCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAG 4042

Qy       2759 ATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGAC 2818
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4043 ATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGAC 4102

Qy       2819 CCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGC 2878
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4103 CCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGC 4162

Qy       2879 AGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCT 2938
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4163 AGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCT 4222

Qy       2939 AGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATC 2998
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4223 AGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATC 4282

Qy       2999 GTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGG 3058
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4283 GTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGG 4342

Qy       3059 CGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATC 3118
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4343 CGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATC 4402

Qy       3119 GTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAAT 3178
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4403 GTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAAT 4462

Qy       3179 TCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAG 3238
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4463 TCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAG 4522

Qy       3239 TCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGAT 3298
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  ||
Db       4523 TCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGAAAT 4582

Qy       3299 AATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGG 3358
             |||||||||||||||||||||||||||||||| |||||||||||||||||||||||| |||
Db       4583 AATACCGCGCCACATAGCAGAACTTTAAAAGTACTCATCATTGGAAAACGTTCTTCAGGG 4642

Qy       3359 CGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCA 3418
             |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db       4643 CGAAAACTCTCAAGGATCTTACCACTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCA 4702

Qy       3419 CCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGA 3478
             |||||||||||||| ||||||||| |||||||  ||  ||||||||||||||||||||||
Db       4703 CCCAACTGATCTTCACAATCTTTTACAAACACCAGCAAATCAAAATGAGCAAAAACAGGA 4762

Qy       3479 AGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTC 3538
             |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db       4763 AGGCAAAATGCCGCAAAAAAGGGAATAAGAGCGACACGGAAATGTTGAATACTCATACTC 4822

Qy       3539 TTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATA 3598
             |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db       4823 TTCCTTTTTCAATATTATTGAAGCATTTATCAGGATTATTGTCTCATGAGCGGATACATA 4882

Qy       3599 TTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTG 3658
             |||||||||||||||||||||||||||||    ||||||||||||||||||||||||||||
Db       4883 TTTGAATGTATTTAGAAAAATAAACAAATAAAAATTCCGCGCACATTTCCCCGAAAAGTG 4942

Qy       3659 CCAC 3662
             ||||
Db       4943 CCAC 4946


RESULT 12
ASPGREEN2/c
LOCUS       ASPGREEN2               3633 bp    DNA     linear   SYN 15-MAY-1998
```

```
DEFINITION  Artificial sequences, plasmid vector pGreen-2.
ACCESSION   Y09374
VERSION     Y09374.1  GI:1684627
KEYWORDS    blaTEM gene; lacZ-PhoC fusion gene; multiple cloning site; T3
            promoter; T7 promoter.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Thaller,M.C., Berlutti,F., Schippa,S., Selan,L. and Rossolini,G.M.
  TITLE     Bacterial acid phosphatase gene fusions useful as targets for
            cloning-dependent insertional inactivation
  JOURNAL   Biotechnol. Prog. 14 (2), 241-247 (1998)
  MEDLINE   98215688
   PUBMED   9548775
REFERENCE   2  (bases 1 to 3633)
  AUTHORS   Rossolini,G.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-NOV-1996) G.M. Rossolini, Dip. Biologia
            Moleculare-Sez., Microbiologia,, Univ. di Siena, via Laterina N.8,
            I- 53100 Siena, ITALY
COMMENT     Related sequences: X64444 and X52328.
FEATURES             Location/Qualifiers
     source          1. .3633
                     /organism="synthetic construct"
                     /mol_type="other DNA"
                     /db_xref="taxon:32630"
                     /plasmid="pGreen-2"
     gene            129. .989
                     /gene="blaTEM"
     CDS             129. .989
                     /gene="blaTEM"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAA70549.1"
                     /db_xref="GI:1684628"
                     /translation="MSIQHFRVALIPFFAAFCLPVFAHPETLVKVKDAEDQLGARVGY
                     IELDLNSGKILESFRPEERFPMMSTFKVLLCGAVLSRIDAGQEQLGRRIHYSQNDLVE
                     YSPVTEKHLTDGMTVRELCSAAITMSDNTAANLLLTTIGGPKELTAFLHNMGDHVTRL
                     DRWEPELNEAIPNDERDTTMPVAMATTLRKLLTGELLTLASRQQLIDWMEADKVAGPL
                     LRSALPAGWFIADKSGAGERGSRGIIAALGPDGKPSRIVVIYTTGSQATMDERNRQIA
                     EIGASLIKHW"
     primer_bind     2147. .2162
                     /note="PUC M13/Reverse Sequencing primer recognition site"
     gene            2154. .3104
                     /gene="lacZ-PhoC"
     CDS             2154. .3104
                     /gene="lacZ-PhoC"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAA70550.1"
                     /db_xref="GI:1684629"
                     /translation="MTMITPSAQLTLTKGNKSWSSTAVAAALELVDPPGCRNSISSLS
                     IPSTSRGGPVPNSPYSESYYARSLAVVLQRRDWENPGVTHGAIPAGNDATTKPDLYYL
                     KNEQAIDSLKLLPPPPEVGSIQFLNDQAMYEKGRMLRNTERGKQAQADADLAAGGVAT
                     AFSGAFGYPITEKDSPELYKLLTNMIEDAGDLATRSAKEHYMRIRPFAFYGTETCNTK
                     DQKKLSTNGSYPSGHTSIGWATALVLAEVNPANQDAILERGYQLGQSRVICGYHWQSD
                     VDAARIVGSAAVATLHSDPAFQAQLAKAKQEFAQKSQKDL"
     promoter        2179. .2198
                     /gene="lacZ-PhoC"
                     /note="T3 promoter"
     misc_feature    2214. .2316
                     /gene="lacZ-PhoC"
                     /note="multiple cloning site"
     promoter        2325. .2344
                     /gene="lacZ-PhoC"
                     /note="T7 promoter"
     primer_bind     2354. .2370
                     /gene="lacZ-PhoC"
                     /note="M13-20 sequencing primer recognition site"
     primer_bind     2374. .2390
                     /gene="lacZ-PhoC"
                     /note="M13-40 sequencing primer recognition site"
```

ORIGIN

Query Match 69.7%; Score 2550.6; DB 12; Length 3633;
Best Local Similarity 82.3%; Pred. No. 0;
Matches 3020; Conservative 0; Mismatches 604; Indels 47; Gaps 6;





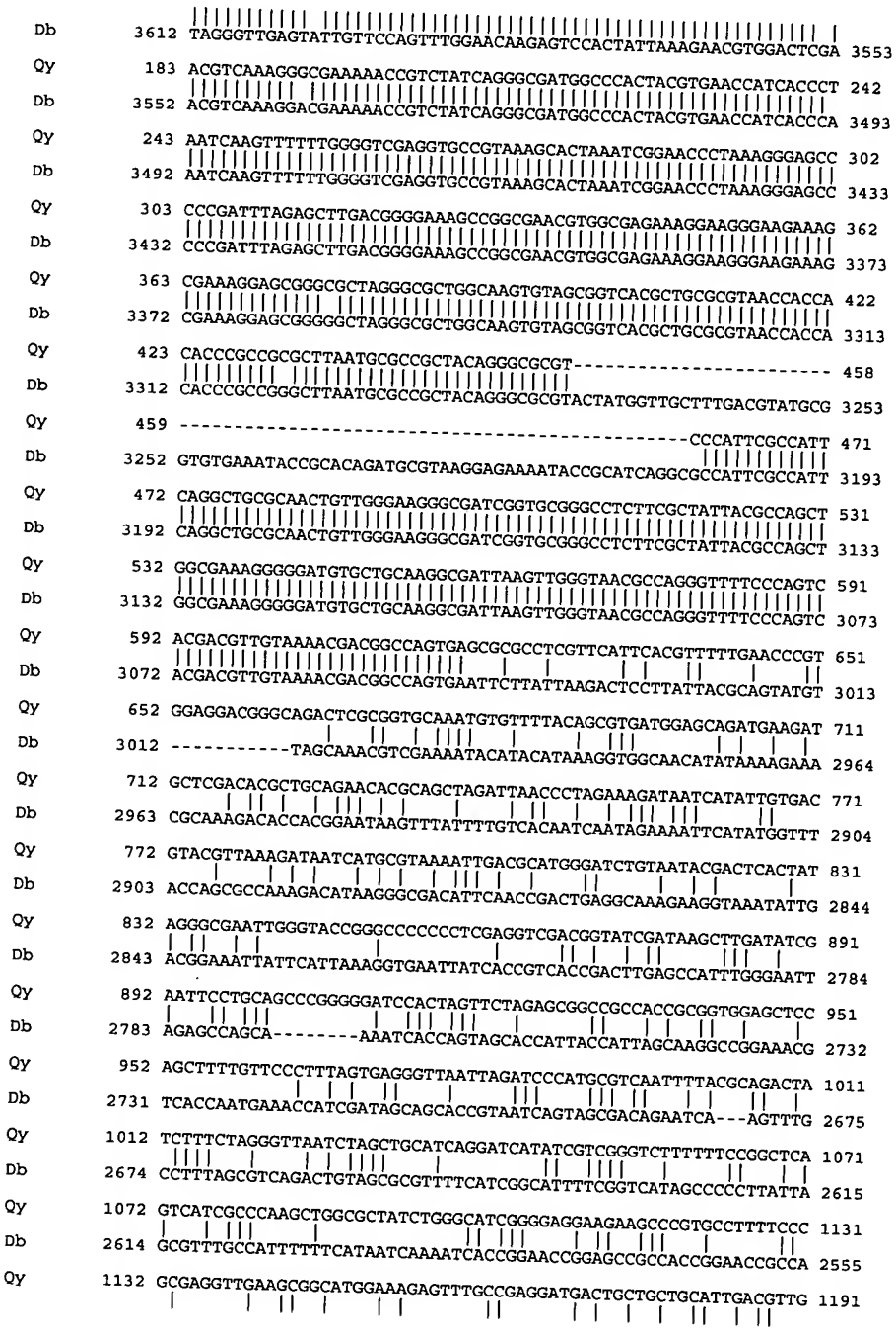
```
Qy       2092 AAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCG 2151
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1571 AAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCG 1512

Qy       2152 CTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGG 2211
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1511 CTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGG 1452

Qy       2212 TAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCAC 2271
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1451 TAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCAC 1392

Qy       2272 TGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTG 2331
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1391 TGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTG 1332

Qy       2332 GCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGT 2391
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1331 GCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGT 1272

Qy       2392 TACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGG 2451
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1271 TACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGG 1212

Qy       2452 TGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCC 2511
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1211 TGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCC 1152

Qy       2512 TTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTT 2571
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1151 TTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTT 1092

Qy       2572 GGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTT 2631
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1091 GGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTT 1032

Qy       2632 TAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAG 2691
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1031 TAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAG 972

Qy       2692 TGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGT 2751
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        971 TGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGT 912

Qy       2752 CGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACC 2811
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        911 CGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACC 852

Qy       2812 GCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGC 2871
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        851 GCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGC 792

Qy       2872 CGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCG 2931
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        791 CGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCG 732

Qy       2932 GGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTAC 2991
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        731 GGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTAC 672

Qy       2992 AGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACG 3051
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        671 AGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACG 612

Qy       3052 ATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCC 3111
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        611 ATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCC 552

Qy       3112 TCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACT 3171
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        551 TCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACT 492

Qy       3172 GCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTC 3231
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        491 GCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTC 432

Qy       3232 AACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAAT 3291
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        431 AACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAAT 372

Qy       3292 ACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTC 3351
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        371 ACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTC 312

Qy       3352 TTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCAC 3411
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        311 TTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCAC 252

Qy       3412 TCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAA 3471
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        251 TCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAA 192

Qy       3472 AACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACT 3531
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        191 AACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACT 132

Qy       3532 CATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGG 3591
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        131 CATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGG 72

Qy       3592 ATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCG 3651
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         71 ATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCG 12

Qy       3652 AAAAGTGCCAC 3662
             ||||||||||||
Db          11 AAAAGTGCCAC 1
```

RESULT 13
CQ768840/c

```
LOCUS       CQ768840               3918 bp    DNA     linear   PAT 04-MAR-2004
DEFINITION  Sequence 59 from Patent WO2004007550.
ACCESSION   CQ768840
VERSION     CQ768840.1  GI:45112122
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Unger,C.M., Zehetmeier,C., Torella,C., Niewoehner,J., Ahrens,B. and
            Beste,G.
  TITLE     Mcam inhibitors
  JOURNAL   Patent: WO 2004007550-A 59 22-JAN-2004;
            Xerion Pharmaceuticals GmbH (DE)
FEATURES             Location/Qualifiers
     source          1. .3918
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="Description of Artificial Sequence: plasmid"
ORIGIN
```

Query Match 68.8%; Score 2517.8; DB 6; Length 3918;
Best Local Similarity 81.8%; Pred. No. 0;
Matches 3052; Conservative 0; Mismatches 537; Indels 142; Gaps 7;

```
Qy          3 AAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTCAT 62
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3732 AAATTGTAAACGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTCAT 3673

Qy         63 TTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGAGA 122
              |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db       3672 TTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGCCCGAGA 3613

Qy        123 TAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTCCA 182
```

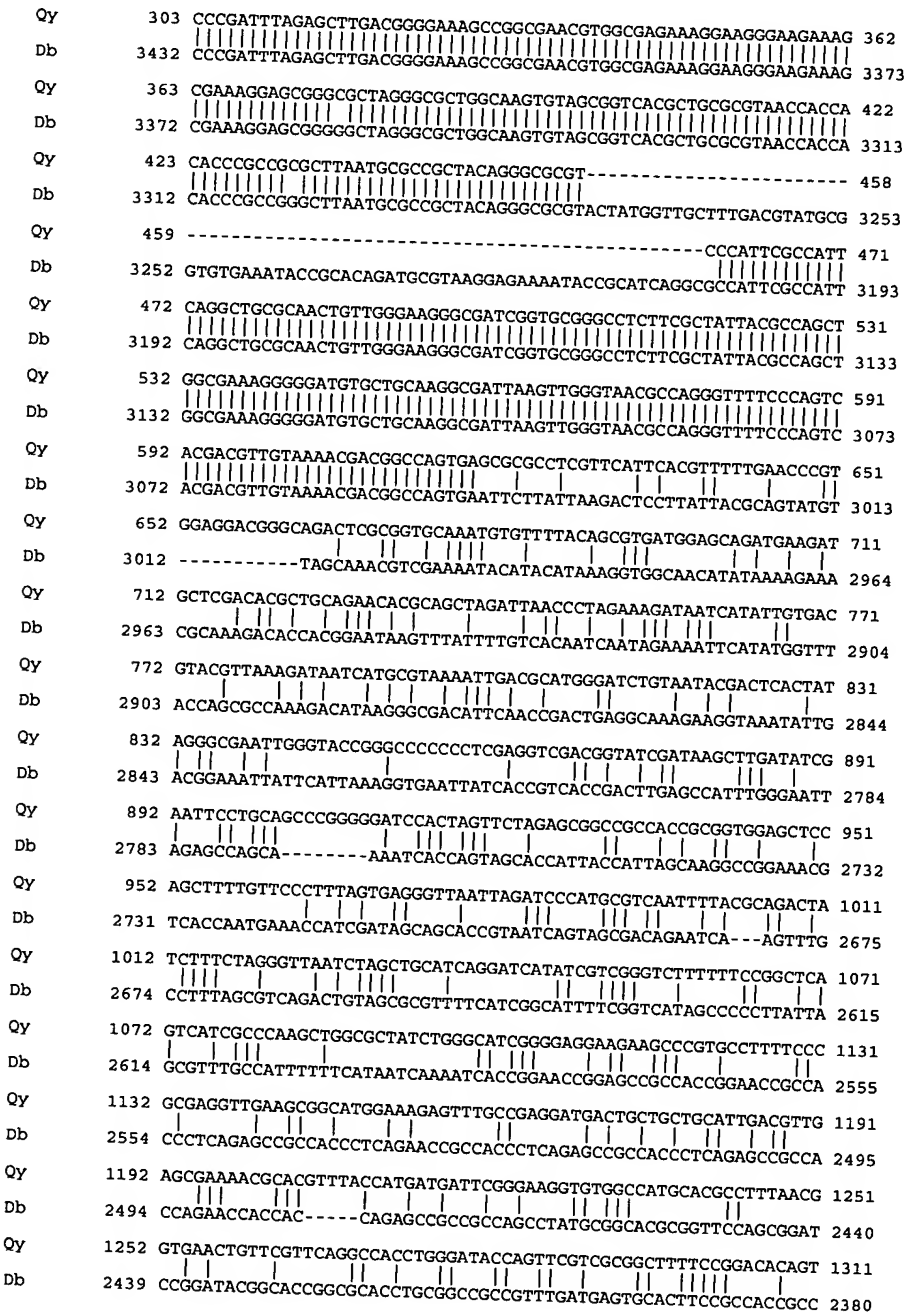

Db 3612 TAGGGTTGAGTATTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTCGA 3553
Qy 183 ACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCCT 242
Db 3552 ACGTCAAAGGACGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCCA 3493
Qy 243 AATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAGCC 302
Db 3492 AATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAGCC 3433
Qy 303 CCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAAAG 362
Db 3432 CCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAAAG 3373
Qy 363 CGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCACCA 422
Db 3372 CGAAAGGAGCGGGGGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCACCA 3313
Qy 423 CACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGT---------- 458
Db 3312 CACCCGCCGGGCTTAATGCGCCGCTACAGGGCGCGTACTATGGTTGCTTTGACGTATGCG 3253
Qy 459 ----------CCCATTCGCCATT 471
Db 3252 GTGTGAAATACCGCACAGATGCGTAAGGAGAAAATACCGCATCAGGCGCCATTCGCCATT 3193
Qy 472 CAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCT 531
Db 3192 CAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCT 3133
Qy 532 GGCGAAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTC 591
Db 3132 GGCGAAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTC 3073
Qy 592 ACGACGTTGTAAAACGACGGCCAGTGAGCGCGCCTCGTTCATTCACGTTTTTGAACCCGT 651
Db 3072 ACGACGTTGTAAAACGACGGCCAGTGAATTCTTATTAAGACTCCTTATTACGCAGTATGT 3013
Qy 652 GGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGCGTGATGGAGCAGATGAAGAT 711
Db 3012 ----------TAGCAAACGTCGAAAATACATACATAAAGGTGGCAACATATAAAAGAAA 2964
Qy 712 GCTCGACACGCTGCAGAACACGCAGCTAGATTAACCCTAGAAAGATAATCATATTGTGAC 771
Db 2963 CGCAAAGACACCACGGAATAAGTTTATTTTGTCACAATCAATAGAAAATTCATATGGTTT 2904
Qy 772 GTACGTTAAAGATAATCATGCGTAAAATTGACGCATGGGATCTGTAATACGACTCACTAT 831
Db 2903 ACCAGCGCCAAAGACATAAGGGCGACATTCAACCGACTGAGGCAAAGAAGGTAAATATTG 2844
Qy 832 AGGGCGAATTGGGTACCGGGCCCCCCCTCGAGGTCGACGGTATCGATAAGCTTGATATCG 891
Db 2843 ACGGAAATTATTCATTAAAGGTGAATTATCACCGTCACCGACTTGAGCCATTTGGGAATT 2784
Qy 892 AATTCCTGCAGCCCGGGGGATCCACTAGTTCTAGAGCGGCCGCCACCGCGGTGGAGCTCC 951
Db 2783 AGAGCCAGCA--------AAATCACCAGTAGCACCATTACCATTAGCAAGGCCGGAAACG 2732
Qy 952 AGCTTTTGTTCCCTTTAGTGAGGGTTAATTAGATCCCATGCGTCAATTTTACGCAGACTA 1011
Db 2731 TCACCAATGAAACCATCGATAGCAGCACCGTAATCAGTAGCGACAGAATCA---AGTTTG 2675
Qy 1012 TCTTTCTAGGGTTAATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCA 1071
Db 2674 CCTTTAGCGTCAGACTGTAGCGCGTTTTCATCGGCATTTTCGGTCATAGCCCCCTTATTA 2615
Qy 1072 GTCATCGCCCAAGCTGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCC 1131
Db 2614 GCGTTTGCCATTTTTTCATAATCAAAATCACCGGAACCGGAGCCGCCACCGGAACCGCCA 2555
Qy 1132 GCGAGGTTGAAGCGGCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTG 1191
Db 2554 CCCTCAGAGCCGCCACCCTCAGAACCGCCACCCTCAGAGCCGCCACCCTCAGAGCCGCCA 2495
Qy 1192 AGCGAAAACGCACGTTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACG 1251
Db 2494 CCAGAACCACCAC-----CAGAGCCGCCGCCAGCCTATGCGGCACGCGGTTCCAGCGGAT 2440
Qy 1252 GTGAACTGTTCGTTCAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGT 1311
Db 2439 CCGGATACGGCACCGGCGCACCTGCGGCCGCCGTTTGATGAGTGCACTTCCGCCACCGCC 2380
Qy 1312 TCCGGATGGTCAGCCCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACT 1371
Db 2379 AGAGCCACCTCCGCCTGAACCGCCTCCACCTGTCGA---------- 2344
Qy 1372 GCCGTGCCGGTGTGCAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGG 1431
Db 2343 -------CTGGACCTGGGCCATGGCCGGCTGGGCCGCATAGAAAGGAACAACTAAAGGAA 2291
Qy 1432 ACGGGTATCCTGGCTGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCG 1491
Db 2290 TTGCGAATAATAATTTTTTCACGTTGAAAATCTCCAAAAAAAA----------GGCT 2244
Qy 1492 GGCGCGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCAC 1551
Db 2243 CCAAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCAC 2184
Qy 1552 AATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGT 1611
Db 2183 AATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGT 2124
Qy 1612 GAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTC 1671
Db 2123 GAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTC 2064
Qy 1672 GTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCG 1731
Db 2063 GTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCG 2004
Qy 1732 CTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGT 1791
Db 2003 CTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGT 1944
Qy 1792 ATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAA 1851
Db 1943 ATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAA 1884
Qy 1852 GAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGC 1911
Db 1883 GAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGC 1824
Qy 1912 GTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAG 1971
Db 1823 GTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAG 1764
Qy 1972 GTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGT 2031
Db 1763 GTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGT 1704
Qy 2032 GCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGG 2091
Db 1703 GCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGG 1644
Qy 2092 AAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCG 2151
Db 1643 AAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCG 1584
Qy 2152 CTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGG 2211
Db 1583 CTCCAAGCTGGGCTGTATGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGG 1524
Qy 2212 TAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCAC 2271
Db 1523 TAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCAC 1464

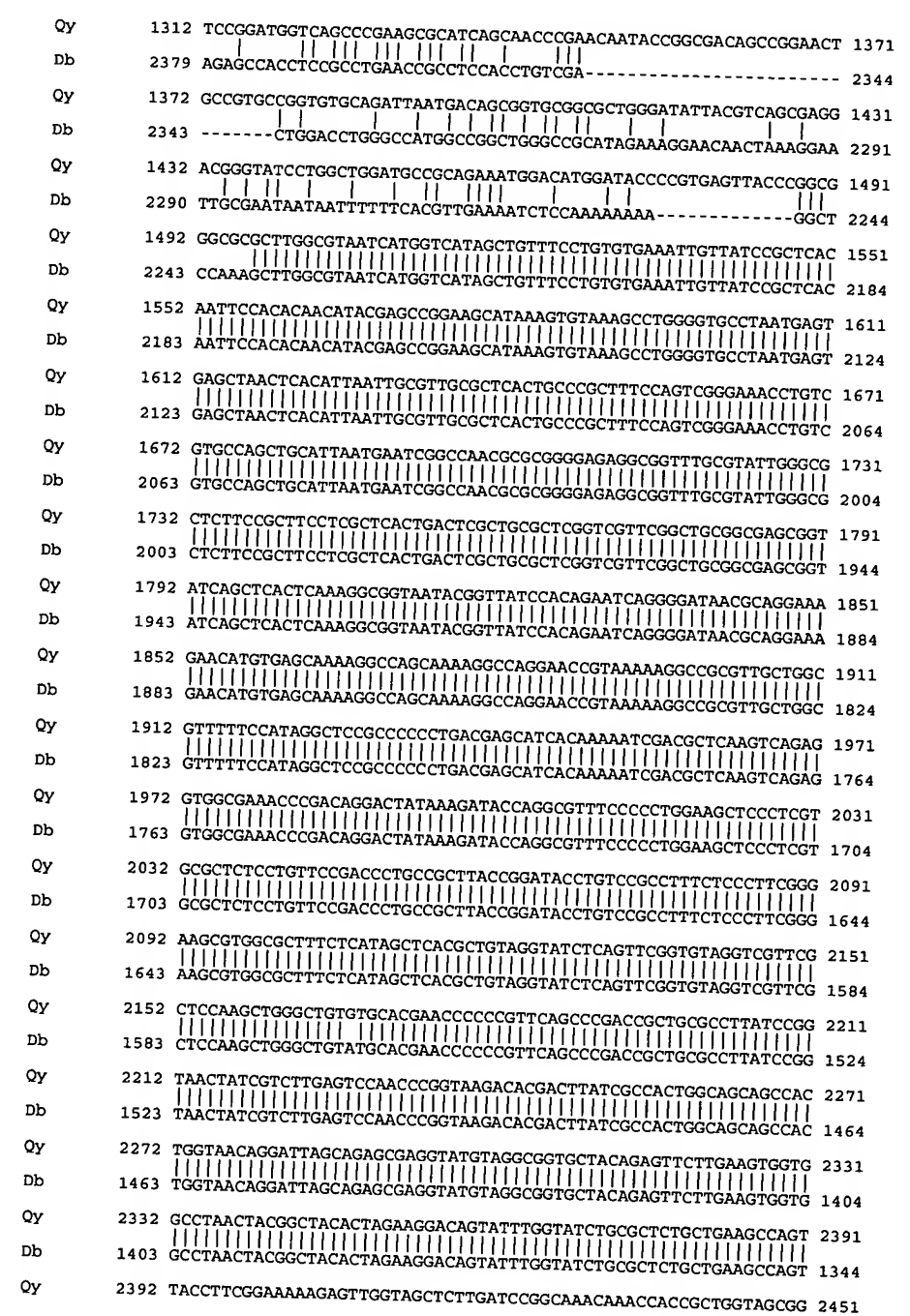
```
Qy       2272 TGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTG 2331
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1463 TGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTG 1404

Qy       2332 GCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGT 2391
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1403 GCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGT 1344

Qy       2392 TACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGG 2451
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1343 TACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGG 1284

Qy       2452 TGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCC 2511
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1283 TGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCC 1224

Qy       2512 TTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTT 2571
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1223 TTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTT 1164

Qy       2572 GGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTT 2631
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1163 GGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTT 1104

Qy       2632 TAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAG 2691
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1103 TAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAG 1044

Qy       2692 TGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGT 2751
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1043 TGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGT 984

Qy       2752 CGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACC 2811
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        983 CGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACC 924

Qy       2812 GCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGC 2871
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        923 GCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGC 864

Qy       2872 CGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCG 2931
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        863 CGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCG 804

Qy       2932 GGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTAC 2991
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        803 GGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTAC 744

Qy       2992 AGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACG 3051
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        743 AGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACG 684

Qy       3052 ATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCC 3111
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        683 ATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCC 624

Qy       3112 TCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACT 3171
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        623 TCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACT 564

Qy       3172 GCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTC 3231
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        563 GCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTC 504

Qy       3232 AACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAAT 3291
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        503 AACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAAT 444

Qy       3292 ACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTC 3351
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        443 ACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTC 384
```

```
Qy       3352 TTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCAC 3411
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        383 TTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCAC 324

Qy       3412 TCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAA 3471
             ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        323 TCGAGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAA 264

Qy       3472 AACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACT 3531
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        263 AACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACT 204

Qy       3532 CATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGG 3591
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 CATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGG 144

Qy       3592 ATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCG 3651
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 ATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCG 84

Qy       3652 AAAAGTGCCAC 3662
             ||||||||||||
Db         83 AAAAGTGCCAC 73
```

```
RESULT 14
CQ802124/c
LOCUS       CQ802124               3918 bp    DNA     linear   PAT 05-MAY-2004
DEFINITION  Sequence 109 from Patent WO2004007717.
ACCESSION   CQ802124
VERSION     CQ802124.1  GI:47058654
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Unger,C.M., Beste,G., Jensen,K.H., Zehetmeier,C., Knauer,R.,
            Guelbenzu,B.L., Torella,C. and Ilag,L.L.
  TITLE     Proteomic screen to identify disease-related biological
            moleculesand inhibitors thereto
  JOURNAL   Patent: WO 2004007717-A 109 22-JAN-2004;
            Xerion Pharmaceuticals AG (DE)
FEATURES             Location/Qualifiers
     source          1. .3918
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="Description of Artificial Sequence: vector pXP10"
ORIGIN

  Query Match             68.8%;  Score 2517.8;  DB 6;  Length 3918;
  Best Local Similarity   81.8%;  Pred. No. 0;
  Matches 3052;  Conservative    0;  Mismatches  537;  Indels  142;  Gaps    7;
```

```
Qy          3 AAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTCAT 62
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3732 AAATTGTAAACGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTCAT 3673

Qy         63 TTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGAGA 122
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db       3672 TTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGCCCGAGA 3613

Qy        123 TAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTCCA 182
              |||||||||| |||||||||||||||||||||||||||||||||||||||||||||| |
Db       3612 TAGGGTTGAGTATTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTCGA 3553

Qy        183 ACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCCT 242
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db       3552 ACGTCAAAGGACGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCCA 3493

Qy        243 AATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAGCC 302
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3492 AATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAGCC 3433
```



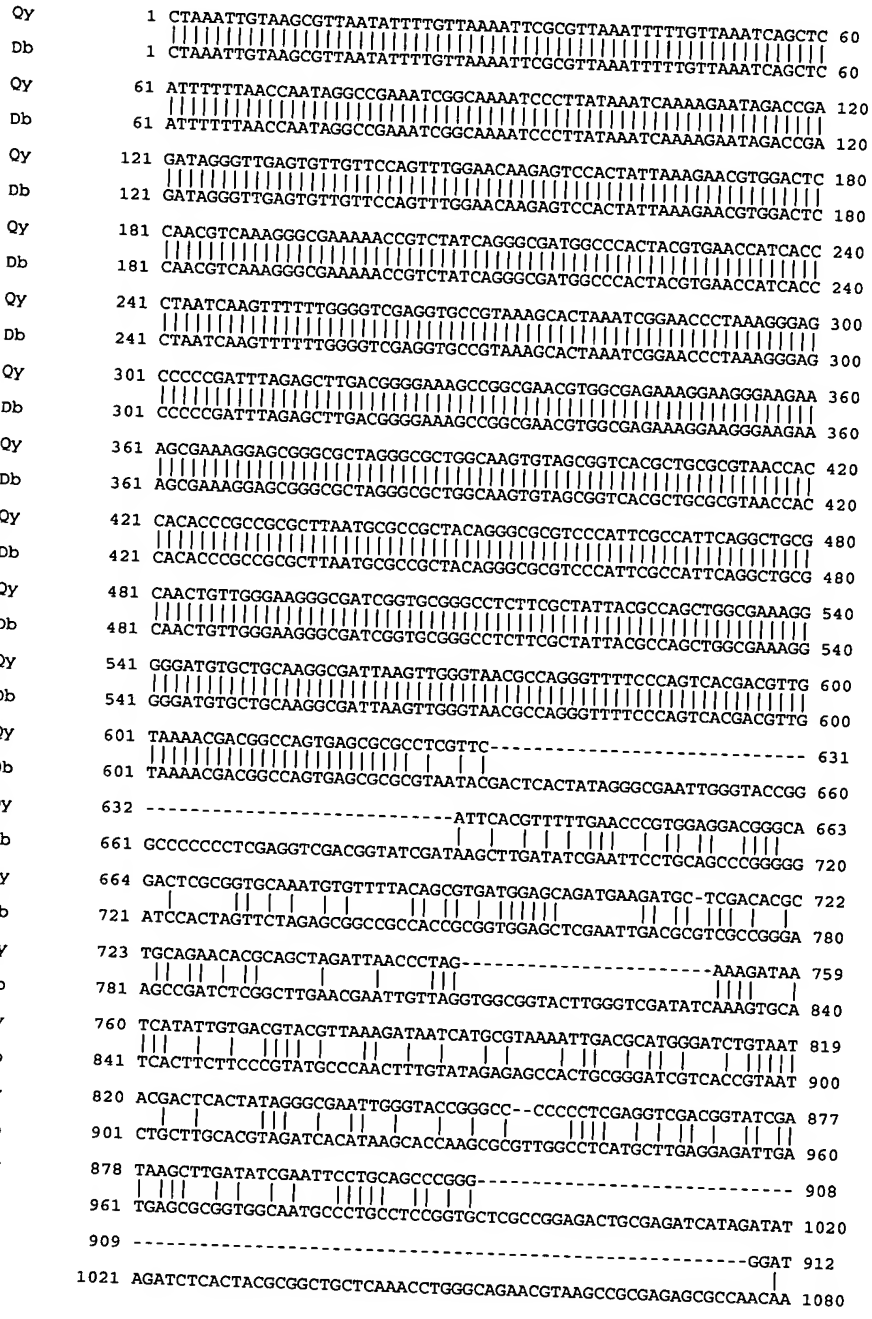


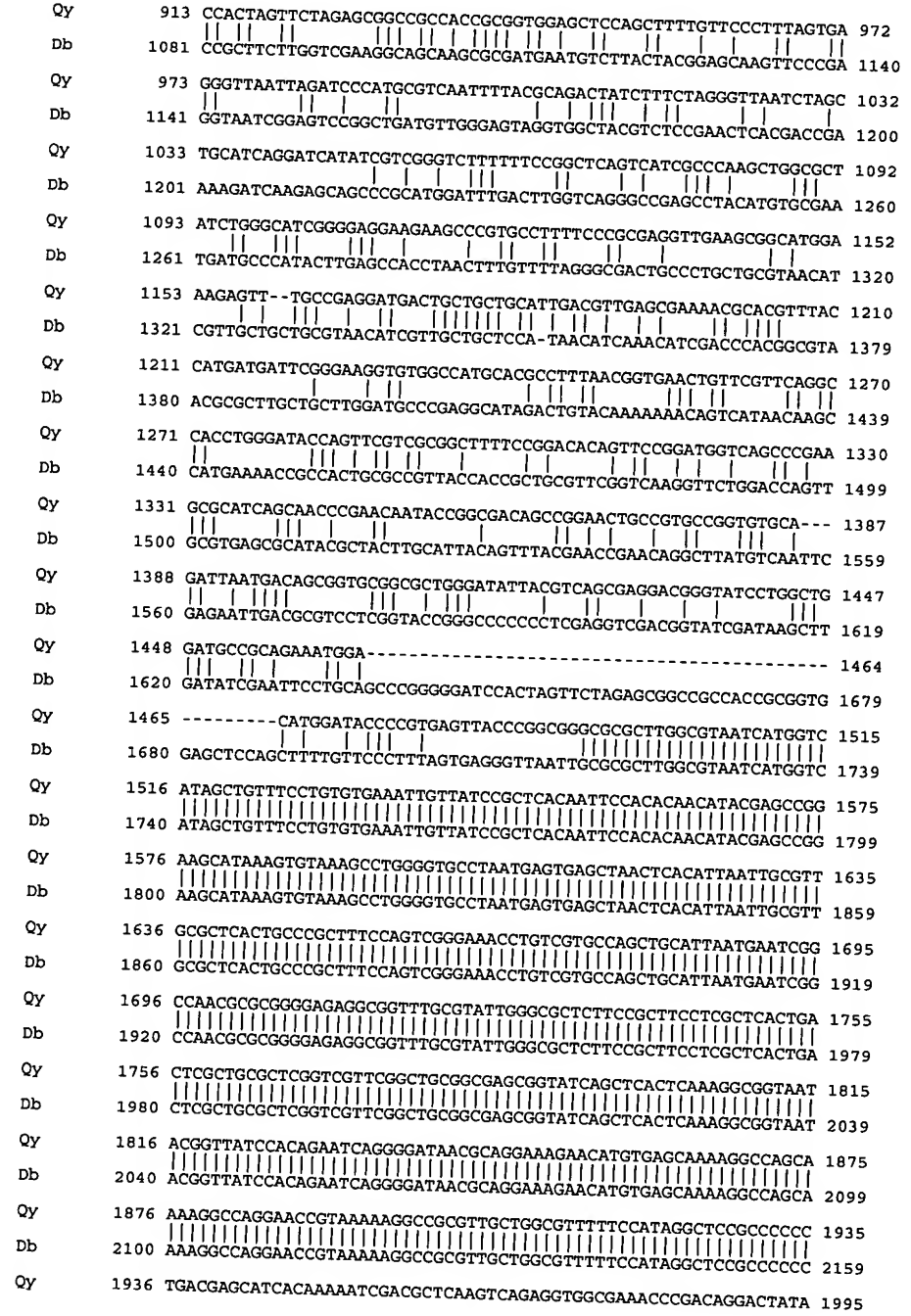
```
Db       1343 TACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGG 1284

Qy       2452 TGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCC 2511
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1283 TGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCC 1224

Qy       2512 TTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTT 2571
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1223 TTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTT 1164

Qy       2572 GGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTT 2631
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1163 GGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTT 1104

Qy       2632 TAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAG 2691
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1103 TAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAG 1044

Qy       2692 TGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGT 2751
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1043 TGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGT 984

Qy       2752 CGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACC 2811
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        983 CGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACC 924

Qy       2812 GCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGC 2871
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        923 GCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGC 864

Qy       2872 CGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCG 2931
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        863 CGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCG 804

Qy       2932 GGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTAC 2991
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        803 GGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTAC 744

Qy       2992 AGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACG 3051
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        743 AGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACG 684

Qy       3052 ATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCC 3111
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        683 ATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCC 624

Qy       3112 TCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACT 3171
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        623 TCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACT 564

Qy       3172 GCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTC 3231
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        563 GCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTC 504

Qy       3232 AACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAAT 3291
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        503 AACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAAT 444

Qy       3292 ACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTC 3351
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        443 ACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTC 384

Qy       3352 TTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCAC 3411
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        383 TTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCAC 324

Qy       3412 TCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAA 3471
             ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        323 TCGAGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAA 264

Qy       3472 AACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACT 3531
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        263 AACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACT 204

Qy       3532 CATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGG 3591
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 CATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGG 144

Qy       3592 ATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCG 3651
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 ATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCG 84

Qy       3652 AAAAGTGCCAC 3662
             ||||||||||| 
Db         83 AAAAGTGCCAC 73
```

RESULT 15
XXU35129
LOCUS XXU35129 3886 bp DNA circular SYN 26-SEP-1995
DEFINITION Plasmid pBSL141 cloning vector, complete sequence.
ACCESSION U35129
VERSION U35129.1 GI:984901
KEYWORDS .
SOURCE Cloning vector pBSL141
ORGANISM Cloning vector pBSL141
artificial sequences; vectors.
REFERENCE 1 (bases 1 to 3886)
AUTHORS Alexeyev,M.F., Shokolenko,I.N. and Croughan,T.P.
TITLE Improved antibiotic-resistance gene cassettes and omega elements for Escherichia coli vector construction and in vitro deletion/insertion mutagenesis
JOURNAL Gene 160 (1), 63-67 (1995)
MEDLINE 95354958
PUBMED 7628718
REFERENCE 2 (bases 1 to 3886)
AUTHORS Hengen,P.N.
TITLE Direct Submission
JOURNAL Submitted (31-AUG-1995) Paul N. Hengen, Laboratory of Mathematical Biology, National Cancer Institute, Frederick, MD 21702-1201, USA
FEATURES Location/Qualifiers
source 1. .3886
/organism="Cloning vector pBSL141"
/mol_type="genomic DNA"
/db_xref="taxon:42702"
/plasmid="pBSL141"
CDS complement(806. .1339)
/codon_start=1
/transl_table=11
/product="gentamicin 3-acetyltransferase"
/protein_id="AAC53614.1"
/db_xref="GI:984902"
/translation="MLRSSNDVTQQGSRPKTKLGGSSMGIIRTCRLGPDQVKSMRAAL
DLFGREFGDVATYSQHQPDSDYLGNLLRSKTFIALAAFDQEAVVGALAAYVLPRFEQP
RSEIYIYDLAVSGEHRRQGIATALINLLKHEANALGAYVIYVQADYGDDPAVALYTKL
GIREEVMHFDIDPSTAT"
CDS complement(2898. .3758)
/EC_number="3.5.2.6"
/function="ampicillin resistance"
/codon_start=1
/transl_table=11
/product="beta-lactamase"
/protein_id="AAC53615.1"
/db_xref="GI:984903"
/translation="MSIQHFRVALIPFFAAFCLPVFAHPETLVKVKDAEDQLGARVGY
IELDLNSGKILESFRPEERFPMMSTFKVLLCGAVLSRIDAGQEQLGRRIHYSQNDLVE
YSPVTEKHLTDGMTVRELCSAAITMSDNTAANLLLTTIGGPKELTAFLHNMGDHVTRL
DRWEPELNEAIPNDERDTTMPVAMATTLRKLLTGELLTLASRQQLIDWMEADKVAGPL
LRSALPAGWFIADKSGAGERGSRGIIAALGPDGKPSRIVVIYTTGSQATMDERNRQIA
EIGASLIKHW"
ORIGIN

Query Match 68.7%; Score 2514.6; DB 12; Length 3886;
Best Local Similarity 80.8%; Pred. No. 0;
Matches 3142; Conservative 0; Mismatches 519; Indels 226; Gaps 9;





```
Db       2160 TGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATA 2219
Qy       1996 AAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCC 2055
Db       2220 AAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCC 2279
Qy       2056 GCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTC 2115
Db       2280 GCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTC 2339
Qy       2116 ACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGA 2175
Db       2340 ACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGA 2399
Qy       2176 ACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCC 2235
Db       2400 ACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCC 2459
Qy       2236 GGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAG 2295
Db       2460 GGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAG 2519
Qy       2296 GTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAG 2355
Db       2520 GTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAG 2579
Qy       2356 GACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAG 2415
Db       2580 GACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAG 2639
Qy       2416 CTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCA 2475
Db       2640 CTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCA 2699
Qy       2476 GATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGA 2535
Db       2700 GATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGA 2759
Qy       2536 CGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGAT 2595
Db       2760 CGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGAT 2819
Qy       2596 CTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGA 2655
Db       2820 CTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGA 2879
Qy       2656 GTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTG 2715
Db       2880 GTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTG 2939
Qy       2716 TCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGA 2775
Db       2940 TCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGA 2999
Qy       2776 GGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCC 2835
Db       3000 GGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCC 3059
Qy       2836 AGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAAC 2895
Db       3060 AGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAAC 3119
Qy       2896 TTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCC 2955
Db       3120 TTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCC 3179
Qy       2956 AGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTC 3015
Db       3180 AGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTC 3239
Qy       3016 GTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCC 3075
Db       3240 GTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCC 3299
Qy       3076 CATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTT 3135
Db       3300 CATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTT 3359
Qy       3136 GGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCC 3195
Db       3360 GGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCC 3419
Qy       3196 ATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTG 3255
Db       3420 ATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTG 3479
Qy       3256 TATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAG 3315
Db       3480 TATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAG 3539
Qy       3316 CAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGAT 3375
Db       3540 CAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGAT 3599
Qy       3376 CTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGC 3435
Db       3600 CTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGC 3659
Qy       3436 ATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAA 3495
Db       3660 ATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAA 3719
Qy       3496 AAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTA 3555
Db       3720 AAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTA 3779
Qy       3556 TTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAA 3615
Db       3780 TTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAA 3839
Qy       3616 AAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCAC 3662
Db       3840 AAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCAC 3886
```

Search completed: January 26, 2005, 08:25:58
Job time : 15663.5 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: January 26, 2005, 02:11:40 ; Search time 11509 Seconds
(without alignments)
11594.581 Million cell updates/sec

Title: US-10-826-523-41
Perfect score: 3662
Sequence: 1 ctaaattgtaagcgttaata.....atttccccgaaaagtgccac 3662

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_htc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1055.8	28.8	1070	1	AJ281552	AJ281552 4A3A-P6F1
c	2	986.4	26.9	1013	4	BM438846	BM438846 IpLvr0015
	3	955.2	26.1	1067	1	AU081137	AU081137 AU081137
	4	924	25.2	1004	1	AJ281480	AJ281480 4A3A-P4G8
	5	923.8	25.2	1089	1	AU081124	AU081124 AU081124
	6	902.8	24.7	1049	9	CL021189	CL021189 CH216-8A1
	7	889.8	24.3	928	7	CO487414	CO487414 GQ0227.B7
c	8	881.4	24.1	925	6	CB686151	CB686151 Bn01b_02o
c	9	870.2	23.8	935	4	BG838279	BG838279 Gc01_10e0
	10	865	23.6	1073	6	CF269652	CF269652 Fcylcold8
	11	857.6	23.4	885	9	CL076016	CL076016 CH216-138
c	12	841	23.0	841	1	AL042026	AL042026 DKFZp434E
c	13	824.6	22.5	854	4	BM438950	BM438950 IpLvr0049
	14	824	22.5	1169	9	AG332951	AG332951 Mus muscu
	15	821.4	22.4	1025	9	CL021193	CL021193 CH216-8A1
	16	820.2	22.4	865	7	CK125894	CK125894 BES182411
	17	819.6	22.4	1249	8	BZ572284	BZ572284 msh2_2572
c	18	814.4	22.2	1126	8	BZ577702	BZ577702 msh2_5533
c	19	809.4	22.1	856	7	CN823189	CN823189 Oa_splbn_
c	20	808	22.1	819	6	CD649375	CD649375 CvGnd0008
	21	807.6	22.1	966	8	BZ570738	BZ570738 msh2_1513
c	22	797.6	21.8	1011	8	BZ576726	BZ576726 msh2_5071
c	23	793	21.7	827	7	CN823902	CN823902 Oa_splbn_
c	24	788.6	21.5	833	6	CB686421	CB686421 Bn01b_04j
	25	786.8	21.5	1163	1	AU081044	AU081044 AU081044
	26	785.4	21.4	1574	8	BZ572566	BZ572566 msh2_2693
	27	777.8	21.2	800	1	AJ281449	AJ281449 4A3A-P4D5
c	28	774.8	21.2	1336	8	BZ575810	BZ575810 msh2_4637
c	29	774.4	21.1	789	6	CD280920	CD280920 G44224.42
	30	772.4	21.1	954	1	AL044364	AL044364 DKFZp434C
c	31	771.4	21.1	832	7	CN822433	CN822433 Oa_splbn_
	32	767.6	21.0	863	7	CF752100	CF752100 TGDR9 Hum
c	33	767.6	21.0	1370	8	BZ571721	BZ571721 msh2_2025
c	34	763.2	20.8	780	5	BQ825693	BQ825693 1030129B0
c	35	757	20.7	759	6	CD279661	CD279661 G43818.35
c	36	755.4	20.6	1463	8	BZ571475	BZ571475 msh2_1906
	37	752.4	20.5	1369	8	BZ579291	BZ579291 msh2_6285
c	38	751.4	20.5	786	7	CN823164	CN823164 Oa_splbn_
	39	750.4	20.5	752	6	CD646006	CD646006 SSPG526R
	40	746.4	20.4	1194	8	BZ578985	BZ578985 msh2_6095
	41	745.4	20.4	844	8	BZ574513	BZ574513 msh2_3706
	42	744.6	20.3	778	5	BP539067	BP539067 BP539067
	43	743.6	20.3	832	4	BG923768	BG923768 602825893
	44	741.4	20.2	758	9	CL422788	CL422788 AE0544 Sa
c	45	741	20.2	741	6	CD279174	CD279174 G44221.83

ALIGNMENTS

RESULT 1
AJ281552
LOCUS AJ281552 1070 bp mRNA linear EST 30-JUN-2000
DEFINITION 4A3A-P6F11-F Anopheles gambiae immune competent 4A3A Anopheles gambiae cDNA clone 4A3A-P6F11, mRNA sequence.
ACCESSION AJ281552
VERSION AJ281552.1 GI:6929432
KEYWORDS EST.
SOURCE Anopheles gambiae (African malaria mosquito)
ORGANISM Anopheles gambiae
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
REFERENCE 1 (bases 1 to 1070)
AUTHORS Dimopoulos,G., Casavant,T.L., Chang,S., Scheetz,T., Roberts,C., Donohue,M., Schultz,J., Benes,V., Bork,P., Ansorge,W., Soares,M.B. and Kafatos,F.C.
TITLE Anopheles gambiae pilot gene discovery project: identification of mosquito innate immunity genes from expressed sequence tags generated from immune-competent cell lines
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (12), 6619-6624 (2000)
MEDLINE 20300950
PUBMED 10841561
COMMENT Contact: Dimopoulos G
Fotis C. Kafatos laboratory
European Molecular Biology Laboratory
Meyerhofstrasse 1, 69117 Heidelberg, Germany.
FEATURES Location/Qualifiers
source 1. .1070
/organism="Anopheles gambiae"
/mol_type="mRNA"
/strain="4A r/r"
/db_xref="taxon:7165"
/clone="4A3A-P6F11"
/cell_line="immune competent 4A3A"
/lab_host="E. coli DH10B"
/clone_lib="Anopheles gambiae immune competent 4A3A"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: EcoRI; Site_2: NotI; sequenced from forward priming site which reads from the 3' end of the cDNA. The 4A3A is a directionally cloned and normalized cDNA library that was constructed from the 4A3A cell line oligo-T primed cDNA according to: Bonaldo, Lennon & Soares (1996) : Normalization and Subtraction: Two approaches To Facilitate Gene Discovery, Genome Research 6, 791-806."
ORIGIN

```
  Query Match           28.8%;  Score 1055.8;  DB 1;  Length 1070;
  Best Local Similarity  99.7%;  Pred. No. 3.8e-299;
  Matches 1068;  Conservative    0;  Mismatches    2;  Indels    1;  Gaps    1;

Qy       1904 TTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCA 1963
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCA 60

Qy       1964 AGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGC 2023
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGC 120

Qy       2024 TCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTC 2083
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTC 180

Qy       2084 CCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAG 2143
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAG 240

Qy       2144 GTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCC 2203
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCC 300

Qy       2204 TTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCA 2263
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCA 360

Qy       2264 GCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTG 2323
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTG 420

Qy       2324 AAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTG 2383
             ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        421 AAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTG 480

Qy       2384 AAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCT 2443
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCT 540

Qy       2444 GGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAA 2503
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAA 600

Qy       2504 GAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAA 2563
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAA 660

Qy       2564 GGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAA 2623
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAA 720

Qy       2624 TGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGC 2683
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGC 780

Qy       2684 TTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGA 2743
             |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        781 TTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTCCATCCATAGTTGCCTGA 840

Qy       2744 CTCCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCA 2803
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CTCCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCA 900

Qy       2804 ATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCC 2863
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 ATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCC 960

Qy       2864 GGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAAT 2923
             |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GGAA-GGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAAT 1019

Qy       2924 TGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAAC 2974
             |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1020 TGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAAC 1070
```

```
RESULT 2
BM438846/c
LOCUS       BM438846               1013 bp    mRNA    linear   EST 31-JAN-2002
DEFINITION  IpLvr00157 Liver cDNA library Ictalurus punctatus cDNA 5', mRNA
            sequence.
ACCESSION   BM438846
VERSION     BM438846.1  GI:18460568
KEYWORDS    EST.
SOURCE      Ictalurus punctatus (channel catfish)
  ORGANISM  Ictalurus punctatus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Siluriformes;
            Ictaluridae; Ictalurus.
REFERENCE   1  (bases 1 to 1013)
  AUTHORS   Feng,J., Kucuktas,H., Kocabas,A., Li,P. and Liu,Z.
  TITLE     Transcriptome of channel catfish (Ictalurus punctatus): initial
            analysis of expressed sequence tags from the liver
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Liu ZJ
            The Fish Molecular Genetics and Biotechnology Laboratory,
            Department of Fisheries and Allied Aquacultures and Program of Cell
            and Molecular Biosciences
            Auburn University
            203 Swingle Hall, Auburn University, Auburn, AL 36849, USA
            Tel: 334 844 4054
            Fax: 334 844 9208
            Email: zliu@acesag.auburn.edu
            Seq primer: M13 Reverse.
FEATURES             Location/Qualifiers
     source          1. .1013
                     /organism="Ictalurus punctatus"
                     /mol_type="mRNA"
                     /db_xref="taxon:7998"
                     /clone_lib="Liver cDNA library"
                     /note="Organ: Liver; Vector: pSport1; Site_1: NotI;
                     Site_2: SalI"
ORIGIN
```

```
  Query Match           26.9%;  Score 986.4;  DB 4;  Length 1013;
  Best Local Similarity  99.5%;  Pred. No. 1.1e-278;
  Matches 1009;  Conservative    1;  Mismatches    2;  Indels    2;  Gaps    2;

Qy       2477 ATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGAC 2536
             ||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||
Db       1013 ATTACGCGCAGAAAAAAAGGATCTCAAGAAGAYCCTTTGATCTTTTCTACGGGGTCTGAC 954

Qy       2537 GCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATC 2596
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        953 GCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATC 894

Qy       2597 TTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAG 2656
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        893 TTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAG 834

Qy       2657 TAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGT 2716
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        833 TAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGT 774

Qy       2717 CTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAG 2776
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        773 CTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAG 714

Qy       2777 GGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCA 2836
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        713 GGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCA 654

Qy       2837 GATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACT 2896
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        653 GATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACT 594
```

```
Qy       2897 TTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCA 2956
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        593 TTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCA 534

Qy       2957 GTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCG 3016
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        533 GTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCG 474

Qy       3017 TTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACA-TGATCCCC 3075
             ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db        473 TTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATTGATCCCC 414

Qy       3076 CATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTT 3135
             |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        413 CATGTTGTGCAAAAAAGCGGTTAGCT-CTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTT 355

Qy       3136 GGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCC 3195
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        354 GGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCC 295

Qy       3196 ATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTG 3255
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        294 ATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTG 235

Qy       3256 TATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAG 3315
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        234 TATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAG 175

Qy       3316 CAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGAT 3375
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        174 CAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGAT 115

Qy       3376 CTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGC 3435
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        114 CTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGC 55

Qy       3436 ATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGC 3489
             |||||| |||||||||||||| |||||||||||||||||||||||||||||||
Db         54 ATCTTTAACTTTCACCAGCGTCTCTGGGTGAGCAAAAACAGGAAGGCAAAATGC 1
```

RESULT 3
AU081137

```
LOCUS       AU081137                1067 bp    mRNA    linear   EST 30-JUL-2002
DEFINITION  AU081137 Oncorhynchus mykiss Kidney infected by infectious
            hematopoietic necrosis virus Oncorhynchus mykiss cDNA clone KI2,
            mRNA sequence.
ACCESSION   AU081137
VERSION     AU081137.1  GI:6431485
KEYWORDS    EST.
SOURCE      Oncorhynchus mykiss (rainbow trout)
  ORGANISM  Oncorhynchus mykiss
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei;
            Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
REFERENCE   1  (bases 1 to 1067)
  AUTHORS   Kono,T., Sakai,M. and LaPatra,S.E.
  TITLE     Expressed Sequence Tag Analysis of Kidney and Gill Tissues from
            Rainbow Trout (Oncorhynchus mykiss) Infected with Infectious
            Hematopoietic Necrosis Virus
  JOURNAL   Mar. Biotechnol. 2 (5), 493-498 (2001)
COMMENT     Contact: Masahiro Sakai
            Faculty of Agriculture
            Miyazaki University
            1-1 nishi gakuenkibanadai, Miyazaki, Miyazaki 889-2192, Japan
            Email: m.sakai@cc.miyazaki-u.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .1067
                     /organism="Oncorhynchus mykiss"
                     /mol_type="mRNA"
                     /db_xref="taxon:8022"
                     /clone="KI2"
                     /tissue_type="kidney"
                     /clone_lib="Oncorhynchus mykiss Kidney infected by
                     infectious hematopoietic necrosis virus"
                     /note="common name:rainbow trout ; infected by infectious
                     hematopoietic necrosis virus"
ORIGIN
```

Query Match 26.1%; Score 955.2; DB 1; Length 1067;
Best Local Similarity 99.0%; Pred. No. 1.7e-269;
Matches 982; Conservative 0; Mismatches 8; Indels 2; Gaps 2;

```
Qy       1494 CGCGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAA 1553
             || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          78 CGAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAA 137

Qy       1554 TTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGA 1613
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         138 TTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGA 197

Qy       1614 GCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGT 1673
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         198 GCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGT 257

Qy       1674 GCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCT 1733
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         258 GCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCT 317

Qy       1734 CTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTAT 1793
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         318 CTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTAT 377

Qy       1794 CAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGA 1853
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         378 CAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGA 437

Qy       1854 ACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGT 1913
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         438 ACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGT 497

Qy       1914 TTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGT 1973
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         498 VVTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGT 557

Qy       1974 GGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGC 2033
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         558 GGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGC 617

Qy       2034 GCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAA 2093
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         618 GCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAA 677

Qy       2094 GCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCT 2153
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         678 GCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCT 737

Qy       2154 CCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTA 2213
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         738 CCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTA 797

Qy       2214 ACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTG 2273
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         798 ACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTG 857

Qy       2274 GTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGC 2333
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         858 GTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGC 917

Qy       2334 CTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTA 2393
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         918 CTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTA 977

Qy       2394 CCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTG 2453
             |||||||||||||||||||||||||||||||||||||||||||||| || ||||||||||||
```

```
Db        978 CCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAA-CAACGCTGGTAGCGTGT 1036

Qy       2454 GTTTTTTTGTTTGCAAGCAGCAGATTACGCGC 2485
              ||||||||||||| ||||||||||||||||||
Db       1037 TTTTTTTGTTTGC-AGCAGCAGATTACGCGC 1067


RESULT 4
AJ281480
LOCUS       AJ281480                1004 bp    mRNA    linear   EST 30-JUN-2000
DEFINITION  4A3A-P4G8-F Anopheles gambiae immune competent 4A3A Anopheles
            gambiae cDNA clone 4A3A-P4G8, mRNA sequence.
ACCESSION   AJ281480
VERSION     AJ281480.1  GI:6929360
KEYWORDS    EST.
SOURCE      Anopheles gambiae (African malaria mosquito)
  ORGANISM  Anopheles gambiae
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;
            Anopheles.
REFERENCE   1  (bases 1 to 1004)
  AUTHORS   Dimopoulos,G., Casavant,T.L., Chang,S., Scheetz,T., Roberts,C.,
            Donohue,M., Schultz,J., Benes,V., Bork,P., Ansorge,W., Soares,M.B.
            and Kafatos,F.C.
  TITLE     Anopheles gambiae pilot gene discovery project: identification of
            mosquito innate immunity genes from expressed sequence tags
            generated from immune-competent cell lines
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (12), 6619-6624 (2000)
  MEDLINE   20300950
   PUBMED   10841561
COMMENT     Contact: Dimopoulos G
            Fotis C. Kafatos laboratory
            European Molecular Biology Laboratory
            Meyerhofstrasse 1, 69117 Heidelberg, Germany.
FEATURES             Location/Qualifiers
     source          1. .1004
                     /organism="Anopheles gambiae"
                     /mol_type="mRNA"
                     /strain="4A r/r"
                     /db_xref="taxon:7165"
                     /clone="4A3A-P4G8"
                     /cell_line="immune competent 4A3A"
                     /lab_host="E. coli DH10B"
                     /clone_lib="Anopheles gambiae immune competent 4A3A"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: EcoRI; Site_2: NotI; sequenced from
                     forward priming site which reads from the 3' end of the
                     cDNA. The 4A3A is a directionally cloned and normalized
                     cDNA library that was constructed from the 4A3A cell line
                     oligo-T primed cDNA according to: Bonaldo, Lennon & Soares
                     (1996) : Normalization and Subtraction: Two approaches To
                     Facilitate Gene Discovery, Genome Research 6, 791-806."
ORIGIN

  Query Match           25.2%;  Score 924;  DB 1;  Length 1004;
  Best Local Similarity   99.0%;  Pred. No. 2.7e-260;
  Matches 1001;  Conservative    1;  Mismatches    2;  Indels    7;  Gaps    7;

Qy       1875 AAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCC 1934
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCC 60

Qy       1935 CTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTAT 1994
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTAT 120

Qy       1995 AAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGC 2054
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGC 180

Qy       2055 CGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCT 2114
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCT 240

Qy       2115 CACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACG 2174
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTNTGCACG 300

Qy       2175 AACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACC 2234
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACC 360

Qy       2235 CGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGA 2294
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGA 420

Qy       2295 GGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAA 2354
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAA 480

Qy       2355 GGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTA 2414
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTA 540

Qy       2415 GCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGC 2474
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGC 600

Qy       2475 AGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTG 2534
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTG 660

Qy       2535 ACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGA 2594
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGA 720

Qy       2595 TCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATG 2654
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATG 780

Qy       2655 AGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCT 2714
              |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        781 AGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCA-CTATCTCAGCGATCT 839

Qy       2715 GTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGG 2774
              |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        840 GTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGAT-ACTACGATACGGG 898

Qy       2775 AGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTC 2834
              |||||||||||||:||| ||||||||||||||||||||||||| ||||||||| ||||||
Db        899 AGGGCTTACCATYTGG-CCCAGTGCTGCAATGATACCGCGAGA-CCACGCTCA-CGGCTC 955

Qy       2835 CAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTG 2885
              ||||||| ||||||||||||||||||||||| ||||||||||||||||||
Db        956 CAGATTT-TCAGCAATAAACCAGCCAGCCGGAA-GGCCGAGCGCAGAAGTG 1004


RESULT 5
AU081124
LOCUS       AU081124                1089 bp    mRNA    linear   EST 30-JUL-2002
DEFINITION  AU081124 Oncorhynchus mykiss Kidney infected by infectious
            hematopoietic necrosis virus Oncorhynchus mykiss cDNA clone KG'12,
            mRNA sequence.
ACCESSION   AU081124
VERSION     AU081124.1  GI:6431472
KEYWORDS    EST.
SOURCE      Oncorhynchus mykiss (rainbow trout)
  ORGANISM  Oncorhynchus mykiss
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei;
            Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
REFERENCE   1  (bases 1 to 1089)
  AUTHORS   Kono,T., Sakai,M. and LaPatra,S.E.
  TITLE     Expressed Sequence Tag Analysis of Kidney and Gill Tissues from
            Rainbow Trout (Oncorhynchus mykiss) Infected with Infectious
            Hematopoietic Necrosis Virus
```

```
  JOURNAL   Mar. Biotechnol. 2 (5), 493-498 (2001)
COMMENT     Contact: Masahiro Sakai
            Faculty of Agriculture
            Miyazaki University
            1-1 nishi gakuenkibanadai, Miyazaki, Miyazaki 889-2192, Japan
            Email: m.sakai@cc.miyazaki-u.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .1089
                     /organism="Oncorhynchus mykiss"
                     /mol_type="mRNA"
                     /db_xref="taxon:8022"
                     /clone="KG'12"
                     /tissue_type="kidney"
                     /clone_lib="Oncorhynchus mykiss Kidney infected by
                     infectious hematopoietic necrosis virus"
                     /note="common name:rainbow trout ; infected by infectious
                     hematopoietic necrosis virus"
ORIGIN

  Query Match           25.2%;  Score 923.8;  DB 1;  Length 1089;
  Best Local Similarity   97.2%;  Pred. No. 3.2e-260;
  Matches  983;  Conservative    0;  Mismatches   22;  Indels    6;  Gaps    4;

Qy       1494 CGCGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAA 1553
             || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         79 CGAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAA 138

Qy       1554 TTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGA 1613
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        139 TTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGA 198

Qy       1614 GCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGT 1673
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        199 GCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGT 258

Qy       1674 GCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCT 1733
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        259 GCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCT 318

Qy       1734 CTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTAT 1793
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        319 CTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTAT 378

Qy       1794 CAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGA 1853
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        379 CAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGA 438

Qy       1854 ACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGT 1913
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        439 ACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGT 498

Qy       1914 TTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGT 1973
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        499 TTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGT 558

Qy       1974 GGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGC 2033
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        559 GGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGC 618

Qy       2034 GCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAA 2093
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        619 GCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAA 678

Qy       2094 GCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCT 2153
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        679 GCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCT 738

Qy       2154 CCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTA 2213
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        739 CCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTA 798

Qy       2214 ACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTG 2273
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        799 ACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTG 858

Qy       2274 GTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAG-TGGTGG 2332
             |||||||||||||||||||||||||||||||||||||||||||||||||||||| | ||||
Db        859 GTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTTTGTGG 918

Qy       2333 CCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTT 2392
             |||||||||||||||||||||| ||||||||| |||||||||||||||| ||||||||
Db        919 CCTAACTACGGCTACACTAGAAGAACAGTATTTTGTATCTGCGCTCTGCTGGAGCCAGTT 978

Qy       2393 ACCTTCGGAAAAAGAGTTGGTAGCTCTT-GATCCGGCAAACAAACCA---CCGCTGGTAG 2448
             |||||||||||||||| |||||||||| |||||||||||||||| ||    |   |   |
Db        979 ACCTTCGGAAAAAGAGTGGGTAGCTCTTGGATCCGGCAAACAAAACAACCGCTGGGAAGG 1038

Qy       2449 CGGTGGTTTTTTTGTTTGCAAGCA-GCAGATTACGCGCAGAAAAAAAGGAT 2498
              ||  || |||||||||||||||| ||||||||| |||  ||||| ||||
Db       1039 GGGGTGTCTTTTTGTTTGCAAGCACGCAGATTACCCGCGAAAAAAACGGAT 1089


RESULT 6
CL021189
LOCUS       CL021189               1049 bp    DNA     linear   GSS 31-DEC-2003
DEFINITION  CH216-8A14_RM1.1 CH216 Xenopus tropicalis genomic clone CH216-8A14,
            genomic survey sequence.
ACCESSION   CL021189
VERSION     CL021189.1  GI:40463002
KEYWORDS    GSS.
SOURCE      Xenopus tropicalis (western clawed frog)
  ORGANISM  Xenopus tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus; Silurana.
REFERENCE   1  (bases 1 to 1049)
  AUTHORS   Kremitzki,C., Carter,J., McPherson,J., Warren,W., Graves,T.,
            Mardis,E. and Wilson,R.
  TITLE     A physical map of the xenopus tropicalis genome
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Richard K Wilson
            Genome Sequencing Center
            Washington University School of Medicine
            Email: submissions@watson.wustl.edu
            Insert Length: 175000   Std Error: 0.00
            Seq primer: RM1 TACGACTCACTATAGGGAGA
            Class: BAC ends
            High quality sequence start: 43
            High quality sequence stop: 888.
FEATURES             Location/Qualifiers
     source          1. .1049
                     /organism="Xenopus tropicalis"
                     /mol_type="genomic DNA"
                     /strain="Nigerian frog"
                     /db_xref="taxon:8364"
                     /clone="CH216-8A14"
                     /sex="male"
                     /cell_line="Stock 248 F7A2, inbred N7"
                     /clone_lib="CH216"
                     /note="Vector: pTARBAC2.1; CHORI-216 Xenopus tropicalis
                     BAC library"
ORIGIN

  Query Match           24.7%;  Score 902.8;  DB 9;  Length 1049;
  Best Local Similarity   98.0%;  Pred. No. 4.9e-254;
  Matches  935;  Conservative    0;  Mismatches   17;  Indels    2;  Gaps    2;

Qy       1486 CCGGCGGGCGCGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCC 1545
             | |||  ||  ||||||||||||||||||||||||||||||||||||||||||||||||||
Db         95 CAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCC 154

Qy       1546 GCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTA 1605
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        155 GCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTA 214

Qy       1606 ATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAA 1665
```

```
Db        215 ATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAA 274

Qy       1666 CCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTAT 1725

Db        275 CCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTAT 334

Qy       1726 TGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCG 1785

Db        335 TGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCG 394

Qy       1786 AGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGC 1845

Db        395 AGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGC 454

Qy       1846 AGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTT 1905

Db        455 AGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTT 514

Qy       1906 GCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAG 1965

Db        515 GCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAG 574

Qy       1966 TCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTC 2025

Db        575 TCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTC 634

Qy       2026 CCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCC 2085

Db        635 CCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCC 694

Qy       2086 TTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGT 2145

Db        695 TTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGT 754

Qy       2146 CGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTT 2205

Db        755 CGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTT 814

Qy       2206 ATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGC 2265

Db        815 ATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGC 874

Qy       2266 AGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAA 2325

Db        875 AGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAAGCGGTGCTACAGAGTTCTTGAA 934

Qy       2326 GTGGTGGCCTAACTACGGCTACACTAGAAGGAC-AGTATTTGGTATCTGCGCTCTGCTGA 2384

Db        935 GTGGTGGCCTAACTACGGCTACACTAGAAGAACAAGTATTTGTTATCTGCGCTCTGCTGA 994

Qy       2385 AGCCAGTTACC-TTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACC 2437

Db        995 AACCAGTTACCTTTCGGAAAAAAAGTTGGTAGCTCTTGAATCCGGCAAAAAAAC 1048
```

```
RESULT 7
CO487414
LOCUS       CO487414                928 bp    mRNA    linear   EST 09-JUL-2004
DEFINITION  GQ0227.B7.1_K04 GQ022: ROOT XYLEM - mature trees Picea glauca cDNA
            clone GQ0227.B7_K04 3', mRNA sequence.
ACCESSION   CO487414
VERSION     CO487414.1  GI:50174316
KEYWORDS    EST.
SOURCE      Picea glauca (white spruce)
  ORGANISM  Picea glauca
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Picea.
REFERENCE   1  (bases 1 to 928)
  AUTHORS   Morency,M.-J., Cooke,J., Pavy,N., Parsons,L., Paule,C., Seguin,A.,
            Retzel,E., Butterfield,Y., Barber,S., Yang,G., Stott,J.,
            Siddiqui,A., Holt,R., Marra,M. and MacKay,J.
  TITLE     Arborea EST sequencing in Picea glauca (white spruce)
  JOURNAL   Unpublished (2004)
COMMENT     Contact: John MacKay
            Centre de Recherche en Biologie Forestiere
            Universite Laval
            Pavillon Charles-Eugene Marchand, Quebec, Quebec, CANADA G1K 7P4
            Fax: 418 656 7493
            Email: jmackay@rsvs.ulaval.ca
            Center for Computational Genomics and Bioinformatics (CCGB),
            University of Minnesota, MN_id Identifier: MN5286152 Clone ID:
            GQ0227.B7_K04 Clones available through: John MacKay, Ph. D.
            Professeur adjoint -Assistant professor EMAIL:
            jmackay@rsvs.ulaval.ca Centre de Recherche en Biologie Forestiere
            (Forest Biology Research Center) Universite Laval Quebec, Quebec
            CANADA G1K 7P4
            Plate: 7.B7  row: 04  column: K
            Seq primer: T7 Primer.
FEATURES             Location/Qualifiers
     source          1. .928
                     /organism="Picea glauca"
                     /mol_type="mRNA"
                     /strain="Three trees Unknown"
                     /db_xref="taxon:3330"
                     /clone="GQ0227.B7_K04"
                     /sex="Hermaphrodite"
                     /tissue_type="Differentiating xylem from roots 1 cm in
                     diameter or larger"
                     /dev_stage="Non-lignified xylem tissues from larger roots
                     in early part of growing season (June)."
                     /lab_host="E. coli DH10B cells"
                     /clone_lib="GQ022: ROOT XYLEM - mature trees"
                     /note="Organ: Roots from 9 year old trees measuring
                     approximately 4.5 m tall, and 10 cm in diameter.; Vector:
                     pBluescript II SK (+) XR; Site_1: Eco-RI; Site_2: Xho-I;
                     cDNA was prepared from 5 mg of poly A+ selected RNA and
                     was directionally ligated into the pBluescript II SK (+)
                     XR vector (Stratagene), transformed by electroporation
                     into DH10B cells (In vitrogen) for propagation"
ORIGIN

  Query Match           24.3%;  Score 889.8;  DB 7;  Length 928;
  Best Local Similarity   96.0%;  Pred. No. 3.2e-250;
  Matches  891;  Conservative    0;  Mismatches   37;  Indels    0;  Gaps    0;

Qy       1915 TTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTG 1974

Db          1 TTTTCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTNNNNNNNG 60

Qy       1975 GCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCG 2034

Db         61 GCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCG 120

Qy       2035 CTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAG 2094

Db        121 CTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAG 180

Qy       2095 CGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTC 2154

Db        181 CGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTC 240

Qy       2155 CAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAA 2214

Db        241 CAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAA 300

Qy       2215 CTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGG 2274

Db        301 CTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGG 360

Qy       2275 TAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCC 2334

Db        361 TAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCC 420

Qy       2335 TAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTAC 2394
```

```
Db        421 TAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTAC 480

Qy       2395 CTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGG 2454
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGG 540

Qy       2455 TTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTT 2514
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTT 600

Qy       2515 GATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGT 2574
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGT 660

Qy       2575 CATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAA 2634
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAA 720

Qy       2635 ATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGA 2694
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGA 780

Qy       2695 GGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGT 2754
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGT 840

Qy       2755 GTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCG 2814
             ||    | ||||||  |||||||||||||||||||  |||       ||||||||||||||  |||
Db        841 GTNNNTNACTACGNTACGGGAGGGCTTACCANCTGNNNNNNGTGCTGCAATGATACNGCG 900

Qy       2815 AGACCCACGCTCACCGGCTCCAGATTTA 2842
             ||||||||||           || |
Db        901 AGACCCACGCNNNNNNNNNNNNNTTNA 928
```

```
RESULT 8
CB686151/c
LOCUS       CB686151                925 bp    mRNA    linear   EST 09-APR-2003
DEFINITION  Bn01b_02o08_A
            Bn01b_AAFC_ECORC_transgenic_Brassica_napus_overexpressing_BNCBF17_c
            onstitutively_frost_tolerant Brassica napus cDNA clone Bn01b_02o08,
            mRNA sequence.
ACCESSION   CB686151
VERSION     CB686151.1  GI:29689876
KEYWORDS    EST.
SOURCE      Brassica napus (rape)
  ORGANISM  Brassica napus
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE   1  (bases 1 to 925)
  AUTHORS   Singh,J., Allard,G., Tinker,N., Robert,L., Lacroix,C., De Moors,A.,
            Chagnon,J., Farah,S., Couroux,P. and Hattori,J.
  TITLE     Expressed Sequence Tags from constitutively frost tolerant
            transgenic Brassica napus overexpressing BNCBF17
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Singh,J.A.
            Eastern Cereal and Oilseed Research Centre
            Agriculture and Agri-food Canada
            KW Neatby Bldg., Central Experimental Farm, Ottawa, Ontario, K1A
            0C6, Canada
            Tel: (613) 759-1662
            Fax: (613) 759-1701
            Email: singhja@agr.gc.ca.
FEATURES             Location/Qualifiers
     source          1. .925
                     /organism="Brassica napus"
                     /mol_type="mRNA"
                     /cultivar="Westar"
                     /db_xref="taxon:3708"
                     /clone="Bn01b_02o08"
                     /tissue_type="fourth leaf"
                     /dev_stage="3 weeks seedling grown at room temperature"
                     /clone_lib="Bn01b_AAFC_ECORC_transgenic_Brassica_napus_ove
                     rexpressing_BNCBF17_constitutively_frost_tolerant"
                     /note="Vector: Bluescript SK+/XhoI-EcoRI; Site_1: EcoRI;
                     Site_2: XhoI; Germinated in soil flats and seedlings grown
                     for 3 weeks in a Conviron E-15 cabinet set at 20oC /16 hr
                     light (250 Em-2sec-1) and 16 oC / 8 hr dark. Fourth leaves
                     collected at 9 am and immediately frozen."
ORIGIN
```

```
  Query Match             24.1%;  Score 881.4;  DB 6;  Length 925;
  Best Local Similarity   99.0%;  Pred. No. 9.7e-248;
  Matches  890;  Conservative    6;  Mismatches    2;  Indels    1;  Gaps    1;

Qy       1493 GCGCGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACA 1552
             |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        898 GCGCGCTTGGCGTAATCATGGTCATAGCTGTTT-CTGTGTGAAATTGTTATCCGCTCACA 840

Qy       1553 ATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTG 1612
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        839 ATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTG 780

Qy       1613 AGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCG 1672
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        779 AGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCG 720

Qy       1673 TGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGC 1732
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        719 TGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGC 660

Qy       1733 TCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTA 1792
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        659 TCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTA 600

Qy       1793 TCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAG 1852
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        599 TCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAG 540

Qy       1853 AACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCG 1912
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        539 AACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCG 480

Qy       1913 TTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGG 1972
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        479 TTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGG 420

Qy       1973 TGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTG 2032
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        419 TGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTG 360

Qy       2033 CGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGA 2092
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        359 CGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGA 300

Qy       2093 AGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGC 2152
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        299 AGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGC 240

Qy       2153 TCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGT 2212
             ||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db        239 TCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGASCGCTGCGCCTTATCCGGT 180

Qy       2213 AACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACT 2272
             ||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||
Db        179 AACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGSAGCAGCCACT 120

Qy       2273 GGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGG 2332
             |||||||:|||||:||||||||||||||||||||||||||||||||||:|||||||||||
Db        119 GGTAACARGATTASCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTSAAGTGGTGG 60

Qy       2333 CCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGT 2391
             | |||||||||:|||||||||| ||||||||||||||||||||||||||||||||||||
Db         59 CTTAACTACGGSTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGT 1
```

```
RESULT 9
BG838279/c
LOCUS       BG838279                935 bp    mRNA    linear   EST 25-MAY-2001
DEFINITION  Gc01_10e07_R Gc01_AAFC_ECORC_cold_stressed_Glycine_clandestina
            Glycine clandestina cDNA clone Gc01_10e07, mRNA sequence.
ACCESSION   BG838279
VERSION     BG838279.1  GI:14204601
KEYWORDS    EST.
SOURCE      Glycine clandestina
  ORGANISM  Glycine clandestina
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
            Glycine.
REFERENCE   1  (bases 1 to 935)
  AUTHORS   Singh,J.A., Farah,S., Chapados,J., Couroux,P., De Moors,A.,
            Harris,L.J., Hattori,J.I., Ouellet,T., Robert,L.S., Sprott,D. and
            Tinker,N.A.
  TITLE     Expressed Sequence Tags from Cold-Stressed Glycine clandestina
            Seedlings
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Singh,J.A.
            Eastern Cereal and Oilseed Research Centre
            Agriculture and Agri-food Canada
            KW Neatby Bldg., Central Experimental Farm, Ottawa, Ontario, K1A
            0C6, Canada
            Tel: (613) 759-1662
            Fax: (613) 759-1701
            Email: singhja@agr.gc.ca.
FEATURES             Location/Qualifiers
     source          1. .935
                     /organism="Glycine clandestina"
                     /mol_type="mRNA"
                     /cultivar="1035"
                     /db_xref="taxon:45687"
                     /clone="Gc01_10e07"
                     /tissue_type="Leaves, stem"
                     /clone_lib="Gc01_AAFC_ECORC_cold_stressed_Glycine_clandest
                     ina"
                     /note="Vector: Bluescript SK+/XhoI-EcoRI; Site_1: EcoRI;
                     Site_2: XhoI; Plants incubated at 2 degrees under 12 hours
                     of light/day.  Harvested  after only 2-3 days of cold
                     treatment. cDNA was prepared with the  Uni-Zap cDNA kit
                     from Stratagene.  Eco RI adapters were linked followed  by
                     digest with Xho I/Eco RI and ligated to pBluescript."
ORIGIN

  Query Match           23.8%;  Score 870.2;  DB 4;  Length 935;
  Best Local Similarity   98.1%;  Pred. No. 2e-244;
  Matches  874;  Conservative   14;  Mismatches    2;  Indels    1;  Gaps    1;

Qy       2773 GGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGC 2832
             || ||||||||||||||| ||||||:||||||||||||||||||:|||||||||:|||||
Db        928 GGGGGGCTTACCATCTGCCCCCAGKGCTGCAATGATACCGCGAGMCCCACGCTCMCCGGC 869

Qy       2833 TCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGC 2892
             :||||||:||||||||||||:|||:||||||||||||||||||||||||||||||||||||
Db        868 YCCAGATWTATCAGCAATAAMCCASCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGC 809

Qy       2893 AACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTC 2952
             |:|||||||||||:||||||||||||||||||||||||||||||||||||||||||||:|
Db        808 AWCTTTATCCGCCYCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTYC 749

Qy       2953 GCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTC 3012
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        748 GCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTC 689

Qy       3013 GTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATC 3072
             ||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||
Db        688 GTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCARGGCGAGTTACATGATC 629

Qy       3073 CCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAA 3132
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        628 CCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAA 569

Qy       3133 GTT-GGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCA 3191
             ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        568 GTTGGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCA 509

Qy       3192 TGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAAT 3251
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        508 TGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAAT 449

Qy       3252 AGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCAC 3311
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        448 AGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCAC 389

Qy       3312 ATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAA 3371
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        388 ATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAA 329

Qy       3372 GGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTT 3431
             ||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||:
Db        328 GGATCTTACCGCTGTTGAGATCCAGTTCGATGTAMCCCACTCGTGCACCCAACTGATCTY 269

Qy       3432 CAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCG 3491
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        268 CAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCG 209

Qy       3492 CAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAAT 3551
             ||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||
Db        208 CAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTYCCTTTTTCAAT 149

Qy       3552 ATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTT 3611
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        148 ATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTT 89

Qy       3612 AGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCAC 3662
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         88 AGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCAC 38

RESULT 10
CF269652
LOCUS       CF269652                1073 bp    mRNA    linear   EST 13-AUG-2003
DEFINITION  Fcylcold844 Fragilariopsis cylindrus SMART cDNA library (Clontech)
            Fragilariopsis cylindrus cDNA clone Antarctic 5', mRNA sequence.
ACCESSION   CF269652
VERSION     CF269652.1  GI:33631539
KEYWORDS    EST.
SOURCE      Fragilariopsis cylindrus
  ORGANISM  Fragilariopsis cylindrus
            Eukaryota; stramenopiles; Bacillariophyta; Bacillariophyceae;
            Bacillariophycidae; Bacillariales; Bacillariaceae; Fragilariopsis.
REFERENCE   1  (bases 1 to 1073)
  AUTHORS   Mock,T. and Valentin,K.
  TITLE     EST analysis of freezing tolerance in the Antarctic diatom
            Fragilariopsis cylindrus: Detection of numerous cold adaption
            related genes and gene transfer events
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Mock T
            Biological Oceanography
            Alfred-Wegener-Intitute for Polar and Marine Research
            Am Handelshafen 12, D-27570 Bremerhaven, Germany
            Tel: +49 471 4831 1893
            Fax: +49 471 4831 1425
            Email: tmock@awi-bremerhaven.de
            sequence with unknown function
            PCR PRimers
            FORWARD: 5'lambdaTriplEX2
            BACKWARD: 3'lambdaTriplEx2
            Seq primer: ctcgggaagcgcgccattgtgttggt.
FEATURES             Location/Qualifiers
     source          1. .1073
                     /organism="Fragilariopsis cylindrus"
```

```
                     /mol_type="mRNA"
                     /strain="Antarctic"
                     /db_xref="taxon:186039"
                     /clone="Antarctic"
                     /clone_lib="Fragilariopsis cylindrus SMART cDNA library
                     (Clontech)"
                     /note="Vector: pTriplEX2; total polyA was used for
                     first-strand synthesis with SMART IV oiligos and CDS
                     III/3'PCR primer. Double strand cDNA synthesis was done by
                     LD PCR using the following program: 95oC for 5 min
                     denaturation and subsequent 20 cycles at 95oC (2min) and
                     68oC (6min). After SfiI digestion the cDNA was
                     fractionated with CHROMA Spin-400 columns. These cDNAs
                     were ligated overnight into pTriplEX2 vectors."
ORIGIN

  Query Match             23.6%;  Score 865;  DB 6;  Length 1073;
  Best Local Similarity   96.2%;  Pred. No. 7e-243;
  Matches  927;  Conservative    0;  Mismatches   30;  Indels    7;  Gaps    4;

Qy       1631 GCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGA 1690
                || |   |||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db          1 GCTTCTTGCTCACTGCCCGCTTTCCAGTCCGGAAACCTGTCGTGCCAGCTGCATTAATGA 60

Qy       1691 ATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTC 1750
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTC 120

Qy       1751 ACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCG 1810
                |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        121 ACTGACTCGCTGCGCTCGGTCTTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCG 180

Qy       1811 GTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGC 1870
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGC 240

Qy       1871 CAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGC 1930
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGC 300

Qy       1931 CCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGA 1990
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGA 360

Qy       1991 CTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACC 2050
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACC 420

Qy       2051 CTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCAT 2110
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCAT 480

Qy       2111 AGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTG 2170
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTG 540

Qy       2171 CACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCC 2230
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCC 600

Qy       2231 AACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGA 2290
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGA 660

Qy       2291 GCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCT--AACTACGGCTACA 2348
                ||||||||||||||||||||||||||||||||||||||||||||    |     |||||||
Db        661 GCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCNTAACTACNGGCTACA 720

Qy       2349 CTAGAAGGACA-GTATTTGGTATCT-GCGCTCTGCTGAAGCC---AGTTACCTTCGGAAA 2403
                ||||||| ||| |||||||||||||| |||||||||  |  |     |  | ||||||||
Db        721 CTAGAAGAACANGTATTTGGTATCTNGCGCTCTGCTNGAAGCCAGNTTACCNTTCGGAAA 780

Qy       2404 AAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGT 2463
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 AAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGT 840

Qy       2464 TTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTC 2523
                |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        841 TTGCAAGCAGCAGATTACGCGCAGAAAAAAAAGATCTCAAGAAGATCCTTTGATCTTTTC 900

Qy       2524 TACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATT 2583
                |||||||||||||||||||||||| ||||||||||||||||  ||||| ||||||
Db        901 TACGGGGTCTGACGCTCAGTGGAACCAAAACTCACGTTAAGGGGATTTGGGCATGAGACA 960

Qy       2584 ATCA 2587
                || |
Db        961 ATAA 964


RESULT 11
CL076016
LOCUS       CL076016                885 bp    DNA     linear   GSS 31-DEC-2003
DEFINITION  CH216-138F20_RM1.1 CH216 Xenopus tropicalis genomic clone
            CH216-138F20, genomic survey sequence.
ACCESSION   CL076016
VERSION     CL076016.1  GI:40531929
KEYWORDS    GSS.
SOURCE      Xenopus tropicalis (western clawed frog)
  ORGANISM  Xenopus tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus; Silurana.
REFERENCE   1  (bases 1 to 885)
  AUTHORS   Kremitzki,C., Carter,J., McPherson,J., Warren,W., Graves,T.,
            Mardis,E. and Wilson,R.
  TITLE     A physical map of the xenopus tropicalis genome
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Richard K Wilson
            Genome Sequencing Center
            Washington University School of Medicine
            Email: submissions@watson.wustl.edu
            Insert Length: 175000   Std Error: 0.00
            Seq primer: RM1 TACGACTCACTATAGGGAGA
            Class: BAC ends
            High quality sequence start: 11
            High quality sequence stop: 810.
FEATURES             Location/Qualifiers
     source          1. .885
                     /organism="Xenopus tropicalis"
                     /mol_type="genomic DNA"
                     /strain="Nigerian frog"
                     /db_xref="taxon:8364"
                     /clone="CH216-138F20"
                     /sex="male"
                     /cell_line="Stock 248 F7A2, inbred N7"
                     /clone_lib="CH216"
                     /note="Vector: pTARBAC2.1; CHORI-216 Xenopus tropicalis
                     BAC library"
ORIGIN

  Query Match             23.4%;  Score 857.6;  DB 9;  Length 885;
  Best Local Similarity   99.3%;  Pred. No. 1e-240;
  Matches  871;  Conservative    0;  Mismatches    5;  Indels    1;  Gaps    1;

Qy       1535 AATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCC 1594
                ||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5 AATTTGTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCC 64

Qy       1595 TGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTC 1654
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         65 TGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTC 124

Qy       1655 CAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGC 1714
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        125 CAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGC 184
```

```
Qy       1715 GGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTT 1774
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        185 GGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTT 244

Qy       1775 CGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCA 1834
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        245 CGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCA 304

Qy       1835 GGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAA 1894
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        305 GGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAA 364

Qy       1895 AAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAAT 1954
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        365 AAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAAT 424

Qy       1955 CGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCC 2014
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        425 CGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCC 484

Qy       2015 CCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCC 2074
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        485 CCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCC 544

Qy       2075 GCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGT 2134
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        545 GCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGT 604

Qy       2135 TCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGAC 2194
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        605 TCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGAC 664

Qy       2195 CGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCG 2254
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        665 CGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCG 724

Qy       2255 CCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACA 2314
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        725 CCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACA 784

Qy       2315 GAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGC 2374
             |||||||||||||||||||||||| |||||||||||| ||||||||||||||||||||||
Db        785 GAGTTCTTGAAGTGGTGGCCTAACTACNGCTACACTAGAAGAACAGTATTTGGTATCTGC 844

Qy       2375 GCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTG 2411
             ||||||||||| |||||||||||||||||||||||| |
Db        845 GCTCTGCTGAAG-CAGTTACCTTCGGAAAAAGAGTGG 880
```

```
RESULT 12
AL042026/c
LOCUS       AL042026                841 bp    mRNA    linear   EST 04-SEP-2003
DEFINITION  DKFZp434E111_r1 434 (synonym: htes3) Homo sapiens cDNA clone
            DKFZp434E111 5', mRNA sequence.
ACCESSION   AL042026
VERSION     AL042026.1  GI:5421372
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 841)
  AUTHORS   Poustka,A., Klein,M., Mewes,H.W., Gassenhuber,J. and Wiemann,S.
  TITLE     EST (Poustka, et al.)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: MIPS
            MIPS
            Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
            This is the 5' sequence of the clone insert
            Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz- heidelberg.de;
            sequenced by DKFZ (German Cancer Research Center,
            Heidelberg/Germany) within the cDNA sequencing consortium of the
            German Genome Project.
            No s1 sequence available.
            This clone (DKFZp434E111) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES             Location/Qualifiers
     source          1. .841
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="DKFZp434E111"
                     /tissue_type="testis"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="434 (synonym: htes3)"
                     /note="Vector: pSport1; Site_1: NotI; Site_2: SalI"
ORIGIN
```

```
  Query Match             23.0%;  Score 841;  DB 1;  Length 841;
  Best Local Similarity   100.0%;  Pred. No. 7.8e-236;
  Matches  841;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2788 TGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGC 2847
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         841 TGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGC 782

Qy       2848 AATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTC 2907
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         781 AATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTC 722

Qy       2908 CATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTT 2967
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         721 CATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTT 662

Qy       2968 GCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGC 3027
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         661 GCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGC 602

Qy       3028 TTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAA 3087
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         601 TTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAA 542

Qy       3088 AAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTT 3147
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         541 AAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTT 482

Qy       3148 ATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATG 3207
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         481 ATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATG 422

Qy       3208 CTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACC 3267
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         421 CTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACC 362

Qy       3268 GAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAA 3327
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 GAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAA 302

Qy       3328 AGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTT 3387
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 AGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTT 242

Qy       3388 GAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTT 3447
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 GAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTT 182

Qy       3448 CACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAG 3507
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 CACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAG 122

Qy       3508 GGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTA 3567
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 GGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTA 62
```

```
Qy       3568 TCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAAT 3627
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAAT 2

Qy       3628 A 3628
              |
Db          1 A 1


RESULT 13
BM438950/c
LOCUS       BM438950                854 bp    mRNA    linear   EST 31-JAN-2002
DEFINITION  IpLvr00491 Liver cDNA library Ictalurus punctatus cDNA 5', mRNA
            sequence.
ACCESSION   BM438950
VERSION     BM438950.1  GI:18460672
KEYWORDS    EST.
SOURCE      Ictalurus punctatus (channel catfish)
  ORGANISM  Ictalurus punctatus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Siluriformes;
            Ictaluridae; Ictalurus.
REFERENCE   1  (bases 1 to 854)
  AUTHORS   Feng,J., Kucuktas,H., Kocabas,A., Li,P. and Liu,Z.
  TITLE     Transcriptome of channel catfish (Ictalurus punctatus): initial
            analysis of expressed sequence tags from the liver
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Liu ZJ
            The Fish Molecular Genetics and Biotechnology Laboratory,
            Department of Fisheries and Allied Aquacultures and Program of Cell
            and Molecular Biosciences
            Auburn University
            203 Swingle Hall, Auburn University, Auburn, AL 36849, USA
            Tel: 334 844 4054
            Fax: 334 844 9208
            Email: zliu@acesag.auburn.edu
            Seq primer: M13 Reverse.
FEATURES             Location/Qualifiers
     source          1. .854
                     /organism="Ictalurus punctatus"
                     /mol_type="mRNA"
                     /db_xref="taxon:7998"
                     /clone_lib="Liver cDNA library"
                     /note="Organ: Liver; Vector: pSport1; Site_1: NotI;
                     Site_2: SalI"
ORIGIN

  Query Match             22.5%;  Score 824.6;  DB 4;  Length 854;
  Best Local Similarity   99.3%;  Pred. No. 5.4e-231;
  Matches  849;  Conservative    0;  Mismatches    4;  Indels    2;  Gaps    2;

Qy       2636 TCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAG 2695
             ||||||| ||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        854 TCAATCTCAAGCATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAG 795

Qy       2696 GCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTG 2755
             ||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db        794 GCACCTATCTCAGCGATCTGTCTATCTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTG 735

Qy       2756 TAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGA 2815
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        734 TAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGA 675

Qy       2816 GACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAG 2875
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        674 GACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAG 615

Qy       2876 CGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAA 2935
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        614 CGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAA 555

Qy       2936 GCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGC 2995
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        554 GCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGC 495

Qy       2996 ATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCA 3055
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        494 ATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCA 435

Qy       3056 AGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCG 3115
             |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        434 AGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCT-CTTCGGTCCTCCG 376

Qy       3116 ATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCAT 3175
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        375 ATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCAT 316

Qy       3176 AATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACC 3235
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        315 AATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACC 256

Qy       3236 AAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGG 3295
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        255 AAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGG 196

Qy       3296 GATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCG 3355
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        195 GATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCG 136

Qy       3356 GGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAA-CCCACTCG 3414
             |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db        135 GGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAAGCCCACTCG 76

Qy       3415 TGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAAC 3474
             ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db         75 TGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTATCTGGGTGAGCAAAAAC 16

Qy       3475 AGGAAGGCAAAATGC 3489
             |||||||||||||||
Db         15 AGGAAGGCAAAATGC 1


RESULT 14
AG332951
LOCUS       AG332951                1169 bp    DNA     linear   GSS 02-JUN-2004
DEFINITION  Mus musculus molossinus DNA, clone:MSMg01-123F02.T7, genomic survey
            sequence.
ACCESSION   AG332951
VERSION     AG332951.1  GI:47906261
KEYWORDS    GSS.
SOURCE      Mus musculus molossinus
  ORGANISM  Mus musculus molossinus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
  TITLE     BAC end Sequences of Library MSMg01
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1169)
  AUTHORS   Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            (E-mail:hattori@gsc.riken.jp, URL:http://hgp.gsc.riken.go.jp/,
            Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT     Clones are derived from the mouse BAC library MSMg01. For BAC
            library availability, please contact Kuniya Abe (abe@rtc.riken.jp).
            Tsukuba Institude, Bio Resource Center,
            The Institute of Physical and Chemical Research (RIKEN) 3-1-1
            Koyadai, Tsukuba, 305-0074 Japan
            phone: 81-298-36-9189, fax: 81-298-36-9199
            e-mail: abe@rtc.riken.jp
            PRIMERS
            Sequencing : T7
```

```
            LIBRARY
            Vector     : pBACe3.6
            R.Site 1   : EcoRI
            R.Site 2   : EcoRI.
FEATURES              Location/Qualifiers
     source           1. .1169
                      /organism="Mus musculus molossinus"
                      /mol_type="genomic DNA"
                      /sub_species="molossinus"
                      /db_xref="taxon:57486"
                      /clone="MSMg01-123F02.T7"
                      /sex="male"
                      /tissue_type="mixture of kidney and spleen"
                      /clone_lib="MSMg01 Mouse Male BAC Library"
ORIGIN

  Query Match             22.5%;  Score 824;  DB 9;  Length 1169;
  Best Local Similarity   89.3%;  Pred. No. 8.8e-231;
  Matches  922;  Conservative    0;  Mismatches  106;  Indels    5;  Gaps    3;

Qy       1486 CCGGCGGGCGCGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCC 1545
             | |||  ||  |||||||||||||| |||||  |||||||||||||||||||||||||||||
Db        130 CAGGCATGCAAGCTTGGCGTAATCCTGGTCCTAGCTGTTTCCTGTGTGAAATTGTTATCC 189

Qy       1546 GCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTA 1605
             |||  |||||||  | ||||||||||||||||||||||||||||||||||||||||||||
Db        190 CCTCCCAATTCCCCCCAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTA 249

Qy       1606 ATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAA 1665
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAA 309

Qy       1666 CCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTAT 1725
             ||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 CCTGTCCTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTAT 369

Qy       1726 TGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCG 1785
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        370 TGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCG 429

Qy       1786 AGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGC 1845
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        430 AGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGC 489

Qy       1846 AGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTT 1905
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        490 AGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTT 549

Qy       1906 GCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAG 1965
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        550 GCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAG 609

Qy       1966 TCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTC 2025
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        610 TCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTC 669

Qy       2026 CCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCC 2085
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        670 CCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCC 729

Qy       2086 TTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGT 2145
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db        730 CTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGT 789

Qy       2146 CGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTT 2205
             |||||||||| |||||||||||||||||||||||||||| |||||||||| ||| |||||||||
Db        790 CGTTCGCTCCCAGCTGGGCTGTGTGCACGAACCCCTCGTTCAGCCCCACCCCTGCGCCTT 849

Qy       2206 ATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGC 2265
             |||||||||||||| || |    | | || |||| |||| |||||||| || ||| ||||
Db        850 ATCCGGTAACTATCCTCCTTGAGTCCAACCCGTAAAACACCACTTATCCCCCCTGNCAGC 909

Qy       2266 AGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAA 2325
```

```
             | |||||||||||||||||||||||| | | |||||||| | |||||||  ||||| ||||
Db        910 ACCCACTGGTAACAGGATTAGCAGAACCAAGTATGTAGCCCGTGCTAC-CAGTTCCTGAA 968

Qy       2326 GTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAA 2385
             | |  |||| |||||| ||||| ||| ||  ||| |||||||| ||| || | || | ||
Db        969 GGGTGGGCCCAACTACCGCTACCCTAAAAAAACACTATTTGGAATCCGCCCCCTCCGAAA 1028

Qy       2386 GCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGG 2445
              ||  |  | | ||  |||| | || ||| |||||||  ||  |||  ||| | || |||
Db        1029 CCCCTTACCTTCCG--AAAAAAATTTGTACCTCTTGACCCCCCAA--AAAACCCCCCTGT 1084

Qy       2446 TAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGA 2505
             | ||| ||||||||||  ||  |  |  || ||  | | || |||||||    | ||| |
Db        1085 AAACGGGGGTTTTTTTTTGCCCCACCCCAAATCCCCCCCAAAAAAAAATTCCCCCAAAA 1144

Qy       2506 AGATCCTTTGATC 2518
              |  |||  ||
Db        1145 TCTTTTTTTTTTC 1157
```

```
RESULT 15
CL021193
LOCUS       CL021193                1025 bp    DNA     linear   GSS 31-DEC-2003
DEFINITION  CH216-8A16_RM1.1 CH216 Xenopus tropicalis genomic clone CH216-8A16,
            genomic survey sequence.
ACCESSION   CL021193
VERSION     CL021193.1  GI:40463006
KEYWORDS    GSS.
SOURCE      Xenopus tropicalis (western clawed frog)
  ORGANISM  Xenopus tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus; Silurana.
REFERENCE   1  (bases 1 to 1025)
  AUTHORS   Kremitzki,C., Carter,J., McPherson,J., Warren,W., Graves,T.,
            Mardis,E. and Wilson,R.
  TITLE     A physical map of the xenopus tropicalis genome
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Richard K Wilson
            Genome Sequencing Center
            Washington University School of Medicine
            Email: submissions@watson.wustl.edu
            Insert Length: 175000   Std Error: 0.00
            Seq primer: RM1 TACGACTCACTATAGGGAGA
            Class: BAC ends
            High quality sequence start: 72
            High quality sequence stop: 837.
FEATURES              Location/Qualifiers
     source           1. .1025
                      /organism="Xenopus tropicalis"
                      /mol_type="genomic DNA"
                      /strain="Nigerian frog"
                      /db_xref="taxon:8364"
                      /clone="CH216-8A16"
                      /sex="male"
                      /cell_line="Stock 248 F7A2, inbred N7"
                      /clone_lib="CH216"
                      /note="Vector: pTARBAC2.1; CHORI-216 Xenopus tropicalis
                      BAC library"
ORIGIN

  Query Match             22.4%;  Score 821.4;  DB 9;  Length 1025;
  Best Local Similarity   95.4%;  Pred. No. 5e-230;
  Matches  900;  Conservative    0;  Mismatches   36;  Indels    7;  Gaps    5;

Qy       1489 GCGGGCGCGCTTGGCGTAATCATGGTCATAG-CTGTTTCCTGTGTGAAATTGTTATC-CG 1546
              ||  ||  |||||||||||||||||||||||| ||||||||||||||||||||||||  ||
Db         84 GCATGCAATCTTGGCGTAATCATGGTCATAGTCTGTTTCCTGTGTGAAATTGTTATCTCG 143

Qy       1547 CTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAA 1606
             |||||||||||||||||| ||| | |||||||||||||| |||||||||| ||||||||||||
Db        144 CTCACAATTCCACACAACTTACTACCCGGAAGCATAAAATGTAAAGCCGGGGGTGCCTAA 203
```

```
Qy       1607 TGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCC--GCTTTCCAGTCGGGAA 1664
             ||||||| ||||||||||||||||||||||||||||||||||   |||||||| |||||||
Db        204 TGAGTGACCTAACTCACATTAATTGCGTTGCGCTCACTGCCCAGACTTTCCAATCGGGAA 263

Qy       1665 ACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGG--GAGAGGCGGTTTGCG 1722
             |||||||||||||||||||||||||||||||||||| ||||||  ||||| |||||||||
Db        264 ACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACACGCGGGGAGAGAGACGGTTTGCG 323

Qy       1723 TATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCG 1782
             |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        324 TATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGATCGTTCGGCTGCA 383

Qy       1783 GCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAA 1842
             |||||||||||||||||||||||||||||||||| |||||||||| ||||| |||||||
Db        384 GCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACAGTTATCCACAAAATCATGGGATAA 443

Qy       1843 CGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGC 1902
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        444 CGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGC 503

Qy       1903 GTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTC 1962
             ||| |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        504 GTTACTGGCGTTTTTCCATAAGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTC 563

Qy       1963 AAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAG 2022
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        564 AAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAG 623

Qy       2023 CTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCT 2082
             |||||||||||||||||||||||||||||||| |||||||||||||| ||||||||||
Db        624 CTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCACTTACCGGATACCTGACCGCCTTTCT 683

Qy       2083 CCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTA 2142
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        684 CCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTA 743

Qy       2143 GGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGC 2202
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        744 GGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGC 803

Qy       2203 CTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGC 2262
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        804 CTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGC 863

Qy       2263 AGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTT 2322
             |||||||||||||||||||||||||||||||||||||||||||||||| | |||||
Db        864 AGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACATACTTCTT 923

Qy       2323 GAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCT 2382
             |||||| |||||||||||||||||| ||||||  ||| ||||||||||||||||||||
Db        924 GAAGTGTTGGCCTAACTACGGCTAC-CTAGAAAAACACTATTTGGTATCTGCGCTCTGCG 982

Qy       2383 GAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCC 2425
             | | |||||||||||||||||| | | ||||||||||||||||
Db        983 GTAACCAGTTACCTTCGGAAAAAAATTGGGTAGCTCTTGATCC 1025
```

Search completed: January 26, 2005, 12:14:02
Job time : 11517 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: January 25, 2005, 23:20:55 ; Search time 1735.03 Seconds
(without alignments)
11079.575 Million cell updates/sec

Title: US-10-826-523-41
Perfect score: 3662
Sequence: 1 ctaaattgtaagcgttaata.....atttccccgaaaagtgccac 3662

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_23Sep04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
	1	3648.4	99.6	3661	10	ACA55354	Aca55354 Transform
c	2	2665.4	72.8	3637	11	ADM68463	Adm68463 Variant g
c	3	2665.4	72.8	3637	12	ADP26616	Adp26616 Green flu
c	4	2664.8	72.8	3637	8	ABX14571	Abx14571 Plasmid p
c	5	2664.8	72.8	3637	11	ADM68448	Adm68448 Cycle 3 g
c	6	2664.8	72.8	3637	12	ADP26601	Adp26601 Green flu
c	7	2656.4	72.5	3637	8	ABX14570	Abx14570 Plasmid p
c	8	2656.4	72.5	3637	11	ADM68462	Adm68462 Wild type
c	9	2656.4	72.5	3637	12	ADP26615	Adp26615 Green flu
c	10	2647.8	72.3	3637	10	ADI26329	Adi26329 Novel end
	11	2635.8	72.0	3681	2	AAQ13578	Aaq13578 Plasmid p
	12	2622	71.6	5314	2	AAQ13576	Aaq13576 Plasmid p
	13	2620.6	71.6	3699	2	AAV14340	Aav14340 Plasmid p
	14	2609.4	71.3	4001	6	AAD27066	Aad27066 Plasmid p
	15	2484.6	67.8	3774	6	AAD27062	Aad27062 Plasmid p
	16	2474.2	67.6	5277	3	AAA88110	Aaa88110 Plamsid p
	17	2464.6	67.3	3908	6	ABT08165	Abt08165 Recombina
	18	2451.8	67.0	3928	8	ABT14478	Abt14478 HCV envel
	19	2451.8	67.0	3928	8	AAD50628	Aad50628 pBKS-E2sH
	20	2451.8	67.0	3928	10	ABQ84236	Abq84236 Vector pB
	21	2448.8	66.9	3956	2	AAV64258	Aav64258 Plasmid p
	22	2446.8	66.8	3715	6	AAD27064	Aad27064 Plasmid T
	23	2446.8	66.8	3927	6	ABT08166	Abt08166 Recombina
	24	2314.8	63.2	4088	2	AAV64255	Aav64255 Plasmid p
	25	2305.2	62.9	4102	2	AAV64257	Aav64257 Plasmid p
	26	2279.2	62.2	4119	6	AAD40772	Aad40772 Plasmid p
	27	2271.2	62.0	4059	12	ADN97131	Adn97131 Mammalian
	28	2230.6	60.9	5250	10	AAD49957	Aad49957 TAR (Tran
	29	2224.6	60.7	2958	2	AAZ22250	Aaz22250 Nucleotid
	30	2216.2	60.5	4205	4	AAD09088	Aad09088 pEAG658 p
	31	2216.2	60.5	4205	6	AAK99713	Aak99713 DNA of pl
c	32	2215.4	60.5	4172	2	AAX19901	Aax19901 Plasmid p
	33	2213	60.4	4205	4	AAD09087	Aad09087 pEAG657 p
	34	2213	60.4	4205	6	AAK99700	Aak99700 DNA of pl
	35	2202.6	60.1	4454	4	AAS06386	Aas06386 Vector pG
	36	2172.8	59.3	7102	2	AAX31836	Aax31836 Nucleotid
	37	2172.8	59.3	7102	6	ABK13259	Abk13259 Yeast pla
	38	2172.8	59.3	7333	2	AAX31837	Aax31837 Nucleotid
	39	2172.8	59.3	7333	6	ABK13260	Abk13260 Yeast pla
c	40	2172.8	59.3	7803	6	ABT08177	Abt08177 Recombina
c	41	2172.8	59.3	8167	6	ABT08178	Abt08178 Recombina
	42	2172.6	59.3	10600	11	ADM68432	Adm68432 Celery CE
	43	2172.6	59.3	10600	12	ADP26585	Adp26585 TMV CEL I
	44	2172.6	59.3	10624	11	ADM68433	Adm68433 Celery CE
	45	2172.6	59.3	10624	12	ADP26586	Adp26586 TMV CEL I

ALIGNMENTS

RESULT 1
ACA55354
ID ACA55354 standard; DNA; 3661 BP.
XX
AC ACA55354;
XX
DT 06-JUN-2003 (first entry)
XX
DE Transformation vector piggyBAC related plasmid pXL-Bac.
XX
KW PiggyBac; transposon; eukaryotic transformation vector; ds;
KW transformed cell; transformed embryo; transgenic; plasmid; cyclic;
KW circular.
XX
OS Synthetic.
XX
PN US2002173634-A1.
XX
PD 21-NOV-2002.
XX
PF 30-OCT-2001; 2001US-00001189.
XX
PR 31-OCT-2000; 2000US-0244677P.
PR 01-NOV-2000; 2000US-0244984P.
XX
PA (FRAS/) FRASER M J.
PA (LIXX/) LI X.
PA (BEAM/) BEAM T.
PA (HUAV/) HUA-VAN A.
XX
PI Fraser MJ, Li X, Beam T, Hua-Van A;
XX
DR WPI; 2003-352597/33.
XX
PT New DNA molecule in the transposon piggyBac, useful for transferring
PT genes into host cells or embryos for transforming the cells of embryos
PT that can be used in making transgenic organisms.
XX
PS Example 3; Fig 3(C2); 151pp; English.
XX
CC The invention describes a DNA molecule comprising at least 163
CC consecutive nucleotide base pairs of the 3' terminal region beginning at
CC the 3' terminal base pair, and at least 125 consecutive nucleotide base
CC pairs of the 5' terminal region beginning at the 5'terminal base pair, of

CC the piggyBac molecule. The region extends from the restriction site SacI
CC to the end of the piggyBac molecule. The DNA molecule in the transposon
CC piggyBac is useful for transferring genes into host cells or embryos for
CC transforming the cells of embryos. The transformed cells or embryos are
CC useful for developing or making transgenic organisms. This sequence
CC represents plasmid used in the creation of minimal sequence eukaryotic
CC transformation vector piggyBac
XX
SQ Sequence 3661 BP; 910 A; 910 C; 957 G; 884 T; 0 U; 0 Other;

Query Match 99.6%; Score 3648.4; DB 10; Length 3661;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3660; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

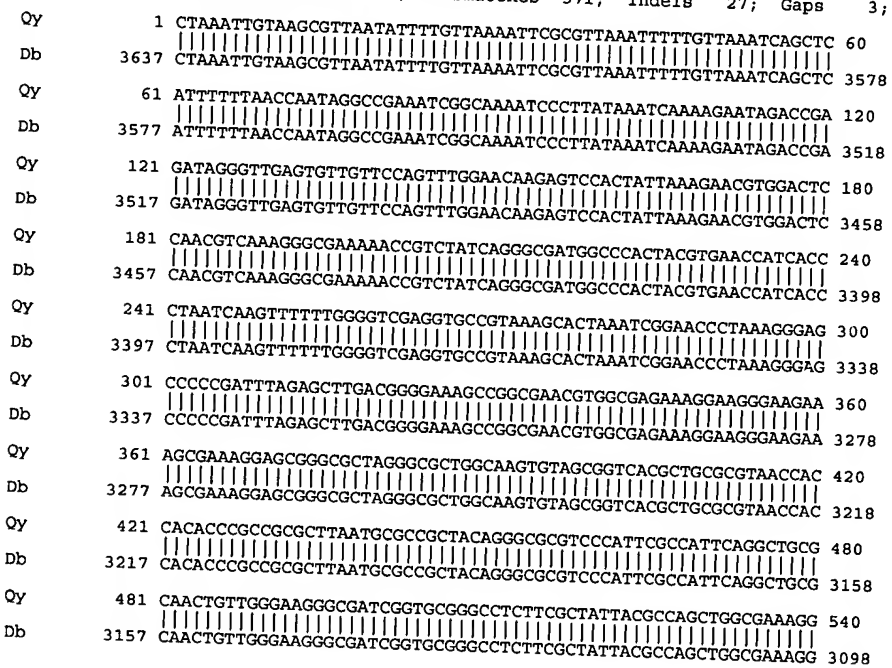
```
Qy       1 CTAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTC 60
Db       1 CTAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTC 60

Qy      61 ATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGA 120
Db      61 ATTTTTTAACCAATAGGCCG-AATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGA 119

Qy     121 GATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTC 180
Db     120 GATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTC 179

Qy     181 CAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACC 240
Db     180 CAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACC 239

Qy     241 CTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAG 300
Db     240 CTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAG 299

Qy     301 CCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAA 360
Db     300 CCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAA 359

Qy     361 AGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCAC 420
Db     360 AGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCAC 419

Qy     421 CACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCCCATTCGCCATTCAGGCTGCG 480
Db     420 CACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCCCATTCGCCATTCAGGCTGCG 479

Qy     481 CAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGG 540
Db     480 CAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGG 539

Qy     541 GGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTG 600
Db     540 GGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTG 599

Qy     601 TAAAACGACGGCCAGTGAGCGCGCCTCGTTCATTCACGTTTTTGAACCCGTGGAGGACGG 660
Db     600 TAAAACGACGGCCAGTGAGCGCGCCTCGTTCATTCACGTTTTTGAACCCGTGGAGGACGG 659

Qy     661 GCAGACTCGCGGTGCAAATGTGTTTTACAGCGTGATGGAGCAGATGAAGATGCTCGACAC 720
Db     660 GCAGACTCGCGGTGCAAATGTGTTTTACAGCGTGATGGAGCAGATGAAGATGCTCGACAC 719

Qy     721 GCTGCAGAACACGCAGCTAGATTAACCCTAGAAAGATAATCATATTGTGACGTACGTTAA 780
Db     720 GCTGCAGAACACGCAGCTAGATTAACCCTAGAAAGATAATCATATTGTGACGTACGTTAA 779

Qy     781 AGATAATCATGCGTAAAATTGACGCATGGGATCTGTAATACGACTCACTATAGGGCGAAT 840
Db     780 AGATAATCATGCGTAAAATTGACGCATGGGATCTGTAATACGACTCACTATAGGGCGAAT 839

Qy     841 TGGGTACCGGGCCCCCCCTCGAGGTCGACGGTATCGATAAGCTTGATATCGAATTCCTGC 900
Db     840 TGGGTACCGGGCCCCCCCTCGAGGTCGACGGTATCGATAAGCTTGATATCGAATTCCTGC 899

Qy     901 AGCCCGGGGGATCCACTAGTTCTAGAGCGGCCGCCACCGCGGTGGAGCTCCAGCTTTTGT 960
Db     900 AGCCCGGGGGATCCACTAGTTCTAGAGCGGCCGCCACCGCGGTGGAGCTCCAGCTTTTGT 959

Qy     961 TCCCTTTAGTGAGGGTTAATTAGATCCCATGCGTCAATTTTACGCAGACTATCTTTCTAG 1020
Db     960 TCCCTTTAGTGAGGGTTAATTAGATCCCATGCGTCAATTTTACGCAGACTATCTTTCTAG 1019

Qy    1021 GGTTAATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCC 1080
Db    1020 GGTTAATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCC 1079

Qy    1081 CAAGCTGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTG 1140
Db    1080 CAAGCTGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTG 1139

Qy    1141 AAGCGGCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAAC 1200
Db    1140 AAGCGGCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAAC 1199

Qy    1201 GCACGTTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGT 1260
Db    1200 GCACGTTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGT 1259

Qy    1261 TCGTTCAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGG 1320
Db    1260 TCGTTCAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGG 1319

Qy    1321 TCAGCCCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCG 1380
Db    1320 TCAGCCCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCG 1379

Qy    1381 GTGTGCAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATC 1440
Db    1380 GTGTGCAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATC 1439

Qy    1441 CTGGCTGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCTT 1500
Db    1440 CTGGCTGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCTT 1499

Qy    1501 GGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACA 1560
Db    1500 GGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACA 1559

Qy    1561 CAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACT 1620
Db    1560 CAACATACCAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACT 1619

Qy    1621 CACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCT 1680
Db    1620 CACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCT 1679

Qy    1681 GCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGC 1740
Db    1680 GCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGC 1739

Qy    1741 TTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCA 1800
Db    1740 TTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCA 1799

Qy    1801 CTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTG 1860
Db    1800 CTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTG 1859

Qy    1861 AGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCA 1920
Db    1860 AGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCA 1919

Qy    1921 TAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAA 1980
Db    1920 TAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAA 1979
```

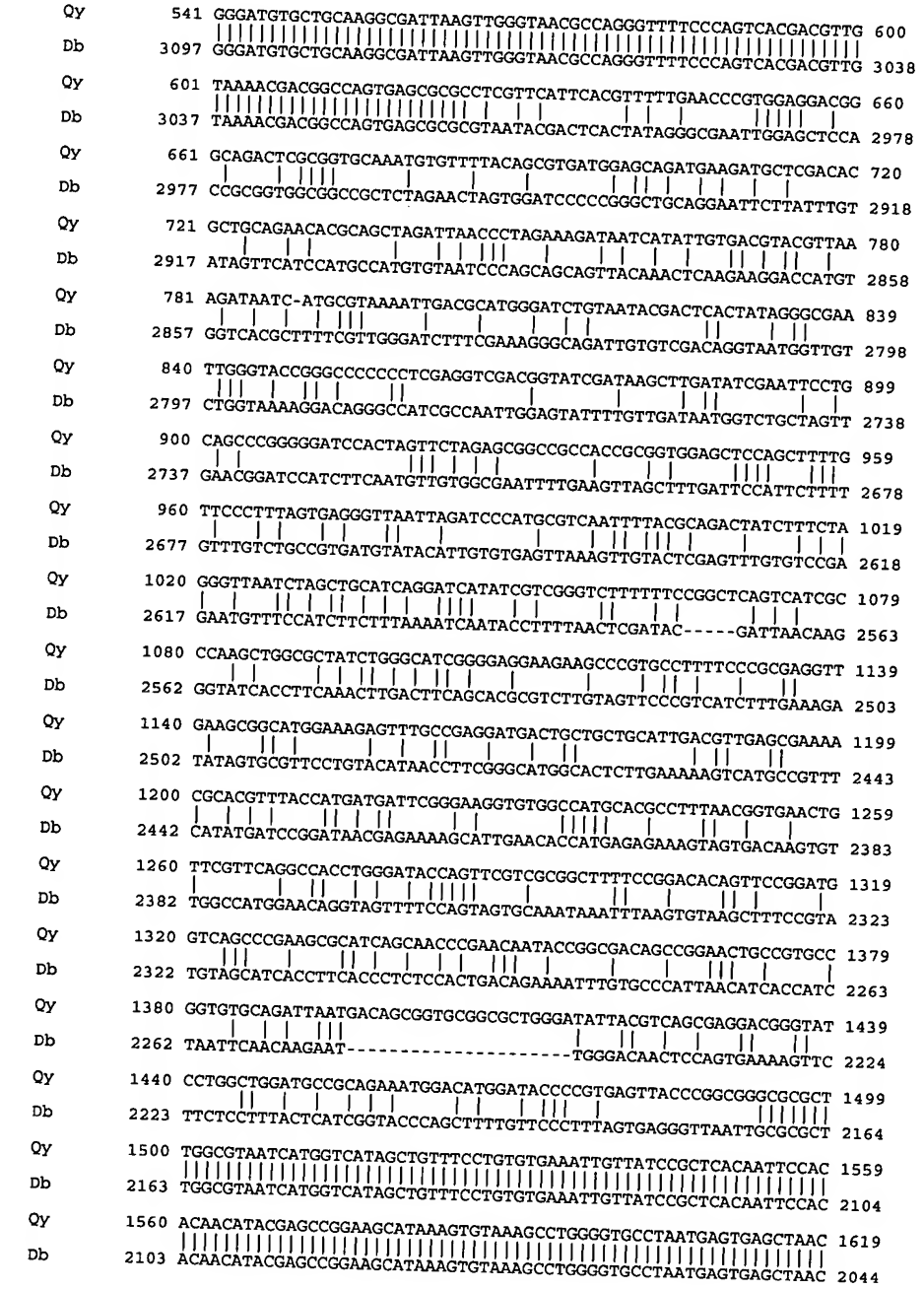
```
Qy       1981 CCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCC 2040
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1980 CCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCC 2039

Qy       2041 TGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGC 2100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2040 TGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGC 2099

Qy       2101 GCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCT 2160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2100 GCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCT 2159

Qy       2161 GGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCG 2220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2160 GGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCG 2219

Qy       2221 TCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAG 2280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2220 TCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAG 2279

Qy       2281 GATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTA 2340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2280 GATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTA 2339

Qy       2341 CGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGG 2400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2340 CGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGG 2399

Qy       2401 AAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTT 2460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2400 AAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTT 2459

Qy       2461 TGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTT 2520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2460 TGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTT 2519

Qy       2521 TTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAG 2580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2520 TTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAG 2579

Qy       2581 ATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAAT 2640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2580 ATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAAT 2639

Qy       2641 CTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACC 2700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2640 CTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACC 2699

Qy       2701 TATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGAT 2760
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2700 TATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGAT 2759

Qy       2761 AACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCC 2820
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2760 AACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCC 2819

Qy       2821 ACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAG 2880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2820 ACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAG 2879

Qy       2881 AAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAG 2940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2880 AAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAG 2939

Qy       2941 AGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGT 3000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2940 AGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGT 2999

Qy       3001 GGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCG 3060
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3000 GGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCG 3059

Qy       3061 AGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGT 3120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3060 AGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGT 3119

Qy       3121 TGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTC 3180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3120 TGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTC 3179

Qy       3181 TCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTC 3240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3180 TCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTC 3239

Qy       3241 ATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAA 3300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3240 ATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAA 3299

Qy       3301 TACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCG 3360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3300 TACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCG 3359

Qy       3361 AAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACC 3420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3360 AAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACC 3419

Qy       3421 CAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAG 3480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3420 CAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAG 3479

Qy       3481 GCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTT 3540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3480 GCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTT 3539

Qy       3541 CCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATT 3600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3540 CCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATT 3599

Qy       3601 TGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCC 3660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3600 TGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCC 3659

Qy       3661 AC 3662
              ||
Db       3660 AC 3661
```

RESULT 2
ADM68463/c
ID ADM68463 standard; DNA; 3637 BP.
XX
AC ADM68463;
XX
DT 03-JUN-2004 (first entry)
XX
DE Variant green fluorescent protein, GFP, gene construct.
XX
KW ds; mismatch endonuclease; endonuclease; gene shuffling technology;
KW single nucleotide polymorphism; cancer susceptibility;
KW sequence variation redistribution; GFP; green fluorescent protein; gene.
XX
OS Aequorea victoria.
OS Synthetic.
XX
PN US2003157682-A1.
XX
PD 21-AUG-2003.
XX
PF 31-JAN-2003; 2003US-00356708.
XX
PR 01-FEB-2002; 2002US-0353722P.
PR 14-MAR-2002; 2002US-00098155.
PR 01-AUG-2002; 2002US-00211079.
XX
PA (PADG/) PADGETT H S.
PA (VAEW/) VAEWHONGS A A.

PA (VOJD/) VOJDANI F S.
PA (SMIT/) SMITH M L.
PA (LIND/) LINDBO J A.
PA (FITZ/) FITZMAURICE W P.
XX
PI Padgett HS, Vaewhongs AA, Vojdani FS, Smith ML, Lindbo JA;
PI Fitzmaurice WP;
XX
DR WPI; 2003-766176/72.
XX
PT Making a mismatch endonuclease, useful in gene shuffling and in detection
PT of single nucleotide polymorphisms, comprises transfecting a host with a
PT recombinant viral vector including a polynucleotide encoding a mismatch
PT endonuclease.
XX
PS Example 5; SEQ ID NO 32; 79pp; English.
XX
CC The invention relates to a method of making a mismatch endonuclease
CC enzyme comprising transfecting a host plant, animal, yeast, fungus or
CC bacterium with a recombinant viral vector that encodes a polynucleotide
CC sequence for a mismatch endonuclease, growing the host so that the
CC polynucleotide is expressed, and extracting the mismatch endonuclease
CC enzyme from the host. The method is useful for making mismatch
CC endonuclease enzymes, for obtaining peptides and polynucleotides with
CC desired functional properties and for detecting mutations. The mismatch
CC endonuclease enzymes are useful in gene shuffling technology for
CC developing new genes, in detecting single nucleotide polymorphisms for
CC e.g. detecting evidence of cancer susceptibility, or in redistributing
CC sequence variations between non-identical polynucleotide sequences. The
CC present sequence represents the variant green fluorescent protein, GFP,
CC gene construct.
XX
SQ Sequence 3637 BP; 942 A; 891 C; 870 G; 934 T; 0 U; 0 Other;

Query Match 72.8%; Score 2665.4; DB 11; Length 3637;
Best Local Similarity 83.7%; Pred. No. 0;
Matches 3065; Conservative 0; Mismatches 571; Indels 27; Gaps 3;





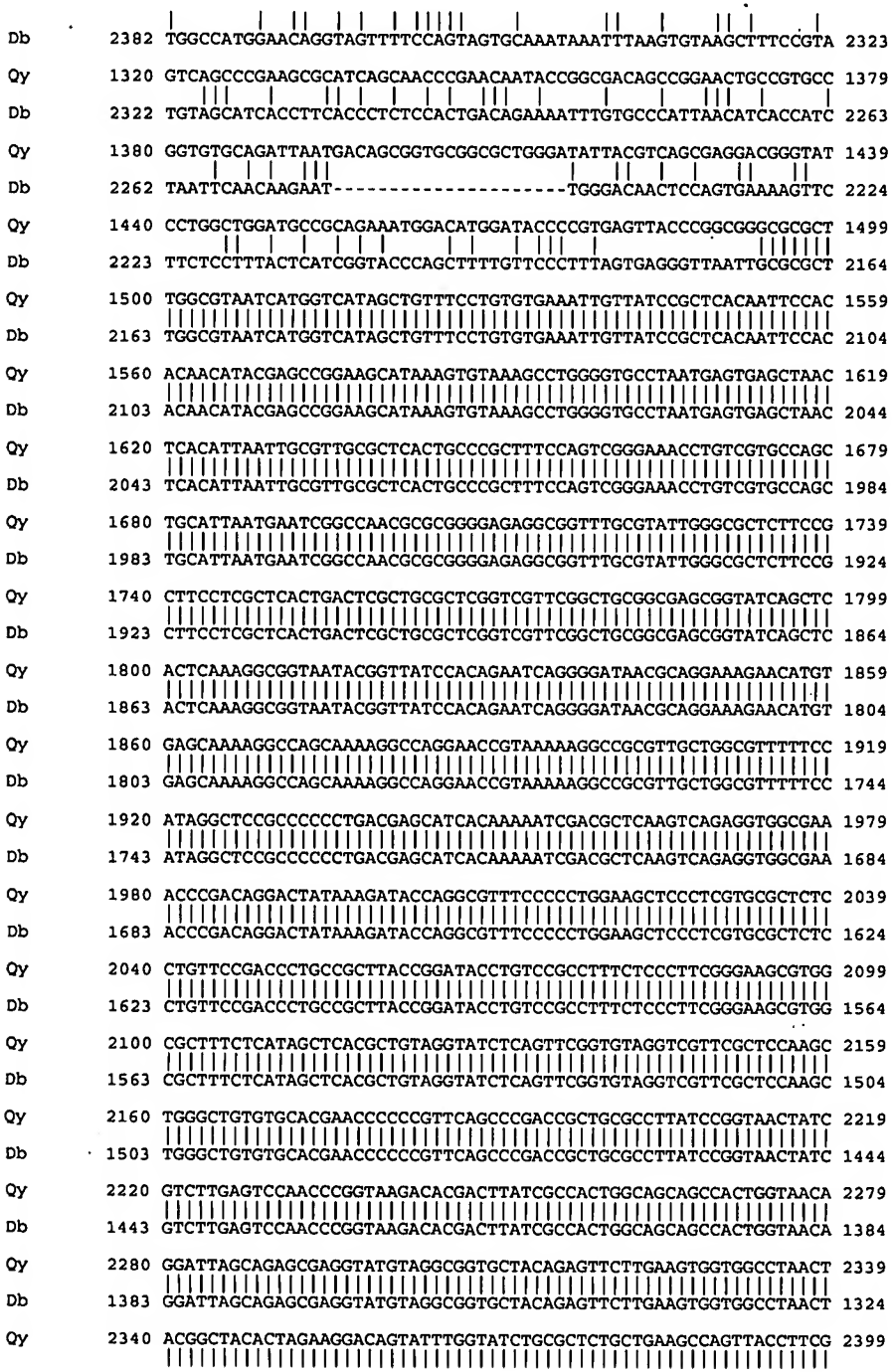
```
Qy       1620 TCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGC 1679
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2043 TCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGC 1984

Qy       1680 TGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCG 1739
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1983 TGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCG 1924

Qy       1740 CTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTC 1799
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1923 CTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTC 1864

Qy       1800 ACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGT 1859
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1863 ACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGT 1804

Qy       1860 GAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCC 1919
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1803 GAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCC 1744

Qy       1920 ATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAA 1979
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1743 ATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAA 1684

Qy       1980 ACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTC 2039
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1683 ACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTC 1624

Qy       2040 CTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGG 2099
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1623 CTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGG 1564

Qy       2100 CGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGC 2159
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1563 CGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGC 1504

Qy       2160 TGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATC 2219
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1503 TGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATC 1444

Qy       2220 GTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACA 2279
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1443 GTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACA 1384

Qy       2280 GGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACT 2339
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1383 GGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACT 1324

Qy       2340 ACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCG 2399
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1323 ACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCG 1264

Qy       2400 GAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTT 2459
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1263 GAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTT 1204

Qy       2460 TTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCT 2519
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1203 TTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCT 1144

Qy       2520 TTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGA 2579
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1143 TTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGA 1084

Qy       2580 GATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAA 2639
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1083 GATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAA 1024

Qy       2640 TCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCAC 2699
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1023 TCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCAC 964

Qy       2700 CTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGA 2759
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        963 CTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGA 904

Qy       2760 TAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACC 2819
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        903 TAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACC 844

Qy       2820 CACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCA 2879
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        843 CACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCA 784

Qy       2880 GAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTA 2939
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        783 GAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTA 724

Qy       2940 GAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCG 2999
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        723 GAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCG 664

Qy       3000 TGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGC 3059
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        663 TGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGC 604

Qy       3060 GAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCG 3119
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        603 GAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCG 544

Qy       3120 TTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATT 3179
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        543 TTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATT 484

Qy       3180 CTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGT 3239
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        483 CTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGT 424

Qy       3240 CATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATA 3299
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        423 CATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATA 364

Qy       3300 ATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGC 3359
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 ATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGC 304

Qy       3360 GAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCAC 3419
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 GAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCAC 244

Qy       3420 CCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAA 3479
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 CCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAA 184

Qy       3480 GGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCT 3539
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 GGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCT 124

Qy       3540 TCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATAT 3599
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 TCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATAT 64

Qy       3600 TTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGC 3659
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          63 TTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGC 4

Qy       3660 CAC 3662
             |||
Db           3 CAC 1
```

RESULT 3
ADP26616/c
ID ADP26616 standard; DNA; 3637 BP.
XX
AC ADP26616;

XX
DT 26-AUG-2004 (first entry)
XX
DE Green fluorescent protein (GFP) plasmid DNA #3.
XX
KW Sequence variation; heteroduplex; transcription; DNA integration;
KW ribozyme expression; gene; ds; green fluorescent protein; GFP.
XX
OS Aequorea victoria.
OS Synthetic.
XX
PN US2004110130-A1.
XX
PD 10-JUN-2004.
XX
PF 25-OCT-2002; 2002US-00280913.
XX
PR 02-FEB-2001; 2001US-0266386P.
PR 14-FEB-2001; 2001US-0268785P.
PR 01-FEB-2002; 2002US-00066390.
PR 08-AUG-2002; 2002US-0402342P.
XX
PA (LARG-) LARGE SCALE BIOLOGY CORP.
XX
PI Padgett HS, Lindbo JA, Fitzmaurice WP;
XX
DR WPI; 2004-440326/41.
XX
PT Redistributing sequence variations between non-identical polynucleotide
PT sequences, useful for generating improved polynucleotide having a desired
PT characteristic, comprises making a heteroduplex and introducing a nick.
XX
PS Example 14; SEQ ID NO 32; 75pp; English.
XX
CC The invention relates to an in vitro method of redistributing sequence
CC variations between non-identical polynucleotide sequences, comprising
CC making a heteroduplex polynucleotide from two non-identical
CC polynucleotides, introducing a nick in the second strand at or near a
CC base pair mismatch site, removing the mismatched base(s) from the
CC mismatch site where the nick occurred and using the first strand as a
CC template to replace the removed base(s) with bases that complement the
CC base(s) in the first strand. The invention also relates to an in vitro
CC method of making a population of sequence variants from a heteroduplex
CC polynucleotide sequence, obtaining a polynucleotide sequence encoding a
CC desired functional property and identifying a reasserted DNA molecule
CC encoding a protein with a desired functional property. The method is
CC useful for generating an improved polynucleotide sequence or a population
CC of improved polynucleotide sequences possessing at least one desired
CC phenotypic characteristic (e.g., promotes transcription of linked
CC polynucleotides), where such polynucleotides are useful for expression
CC from a plant, animal, fungal, yeast, or bacterial expression vector, for
CC integration to form a transgenic plant, animal or microorganism, and for
CC expression of a ribozyme. This sequence represents DNA used in the scope
CC of the invention.
XX
SQ Sequence 3637 BP; 942 A; 891 C; 870 G; 934 T; 0 U; 0 Other;

Query Match 72.8%; Score 2665.4; DB 12; Length 3637;
Best Local Similarity 83.7%; Pred. No. 0;
Matches 3065; Conservative 0; Mismatches 571; Indels 27; Gaps 3;

```
Qy       1 CTAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTC 60
Db    3637 CTAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTC 3578

Qy      61 ATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGA 120
Db    3577 ATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGA 3518

Qy     121 GATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTC 180
Db    3517 GATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTC 3458

Qy     181 CAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACC 240
Db    3457 CAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACC 3398

Qy     241 CTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAG 300
Db    3397 CTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAG 3338

Qy     301 CCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAA 360
Db    3337 CCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAA 3278

Qy     361 AGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCAC 420
Db    3277 AGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCAC 3218

Qy     421 CACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCCCATTCGCCATTCAGGCTGCG 480
Db    3217 CACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCCCATTCGCCATTCAGGCTGCG 3158

Qy     481 CAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGG 540
Db    3157 CAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGG 3098

Qy     541 GGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTG 600
Db    3097 GGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTG 3038

Qy     601 TAAAACGACGGCCAGTGAGCGCGCGTAATACGACTCACTATAGGGCGAATTGGAGCTCCA 660
Db    3037 TAAAACGACGGCCAGTGAGCGCGCGTAATACGACTCACTATAGGGCGAATTGGAGCTCCA 2978

Qy     661 GCAGACTCGCGGTGCAAATGTGTTTTACAGCGTGATGGAGCAGATGAAGATGCTCGACAC 720
Db    2977 CCGCGGTGGCGGCCGCTCTAGAACTAGTGGATCCCCCGGGCTGCAGGAATTCTTATTTGT 2918

Qy     721 GCTGCAGAACACGCAGCTAGATTAACCCTAGAAAGATAATCATATTGTGACGTACGTTAA 780
Db    2917 ATAGTTCATCCATGCCATGTGTAATCCCAGCAGCAGTTACAAACTCAAGAAGGACCATGT 2858

Qy     781 AGATAATC-ATGCGTAAAATTGACGCATGGGATCTGTAATACGACTCACTATAGGGCGAA 839
Db    2857 GGTCACGCTTTTCGTTGGGATCTTTCGAAAGGGCAGATTGTGTCGACAGGTAATGGTTGT 2798

Qy     840 TTGGGTACCGGGCCCCCCCTCGAGGTCGACGGTATCGATAAGCTTGATATCGAATTCCTG 899
Db    2797 CTGGTAAAAGGACAGGGCCATCGCCAATTGGAGTATTTTGTTGATAATGGTCTGCTAGTT 2738

Qy     900 CAGCCCGGGGGATCCACTAGTTCTAGAGCGGCCGCCACCGCGGTGGAGCTCCAGCTTTTG 959
Db    2737 GAACGGATCCATCTTCAATGTTGTGGCGAATTTTGAAGTTAGCTTTGATTCCATTCTTTT 2678

Qy     960 TTCCCTTTAGTGAGGGTTAATTAGATCCCATGCGTCAATTTTACGCAGACTATCTTTCTA 1019
Db    2677 GTTTGTCTGCCGTGATGTATACATTGTGTGAGTTAAAGTTGTACTCGAGTTTGTGTCCGA 2618

Qy    1020 GGGTTAATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGC 1079
Db    2617 GAATGTTTCCATCTTCTTTAAAATCAATACCTTTTAACTCGATAC-----GATTAACAAG 2563

Qy    1080 CCAAGCTGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTT 1139
Db    2562 GGTATCACCTTCAAACTTGACTTCAGCACGCGTCTTGTAGTTCCCGTCATCTTTGAAAGA 2503

Qy    1140 GAAGCGGCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAA 1199
Db    2502 TATAGTGCGTTCCTGTACATAACCTTCGGGCATGGCACTCTTGAAAAAGTCATGCCGTTT 2443

Qy    1200 CGCACGTTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTG 1259
Db    2442 CATATGATCCGGATAACGAGAAAAGCATTGAACACCATGAGAGAAAGTAGTGACAAGTGT 2383

Qy    1260 TTCGTTCAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATG 1319
```


Db 2382 TGGCCATGGAACAGGTAGTTTTCCAGTAGTGCAAATAAATTTAAGTGTAAGCTTTCCGTA 2323
Qy 1320 GTCAGCCCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCC 1379
Db 2322 TGTAGCATCACCTTCACCCTCTCCACTGACAGAAAATTTGTGCCCATTAACATCACCATC 2263
Qy 1380 GGTGTGCAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTAT 1439
Db 2262 TAATTCAACAAGAAT----------TGGGACAACTCCAGTGAAAAGTTC 2224
Qy 1440 CCTGGCTGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCT 1499
Db 2223 TTCTCCTTTACTCATCGGTACCCAGCTTTTGTTCCCTTTAGTGAGGGTTAATTGCGCGCT 2164
Qy 1500 TGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCAC 1559
Db 2163 TGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCAC 2104
Qy 1560 ACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAAC 1619
Db 2103 ACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAAC 2044
Qy 1620 TCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGC 1679
Db 2043 TCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGC 1984
Qy 1680 TGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCG 1739
Db 1983 TGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCG 1924
Qy 1740 CTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTC 1799
Db 1923 CTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTC 1864
Qy 1800 ACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGT 1859
Db 1863 ACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGT 1804
Qy 1860 GAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCC 1919
Db 1803 GAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCC 1744
Qy 1920 ATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAA 1979
Db 1743 ATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAA 1684
Qy 1980 ACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTC 2039
Db 1683 ACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTC 1624
Qy 2040 CTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGG 2099
Db 1623 CTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGG 1564
Qy 2100 CGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGC 2159
Db 1563 CGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGC 1504
Qy 2160 TGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATC 2219
Db 1503 TGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATC 1444
Qy 2220 GTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACA 2279
Db 1443 GTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACA 1384
Qy 2280 GGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACT 2339
Db 1383 GGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACT 1324
Qy 2340 ACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCG 2399
Db 1323 ACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCG 1264
Qy 2400 GAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTT 2459
Db 1263 GAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTT 1204
Qy 2460 TTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCT 2519
Db 1203 TTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCT 1144
Qy 2520 TTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGA 2579
Db 1143 TTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGA 1084
Qy 2580 GATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAA 2639
Db 1083 GATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAA 1024
Qy 2640 TCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCAC 2699
Db 1023 TCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCAC 964
Qy 2700 CTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGA 2759
Db 963 CTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGA 904
Qy 2760 TAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACC 2819
Db 903 TAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACC 844
Qy 2820 CACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCA 2879
Db 843 CACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCA 784
Qy 2880 GAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTA 2939
Db 783 GAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTA 724
Qy 2940 GAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCG 2999
Db 723 GAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCG 664
Qy 3000 TGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGC 3059
Db 663 TGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGC 604
Qy 3060 GAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCG 3119
Db 603 GAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCG 544
Qy 3120 TTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATT 3179
Db 543 TTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATT 484
Qy 3180 CTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGT 3239
Db 483 CTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGT 424
Qy 3240 CATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATA 3299
Db 423 CATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATA 364
Qy 3300 ATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGC 3359
Db 363 ATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGC 304
Qy 3360 GAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCAC 3419
Db 303 GAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCAC 244
Qy 3420 CCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAA 3479
Db 243 CCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAA 184

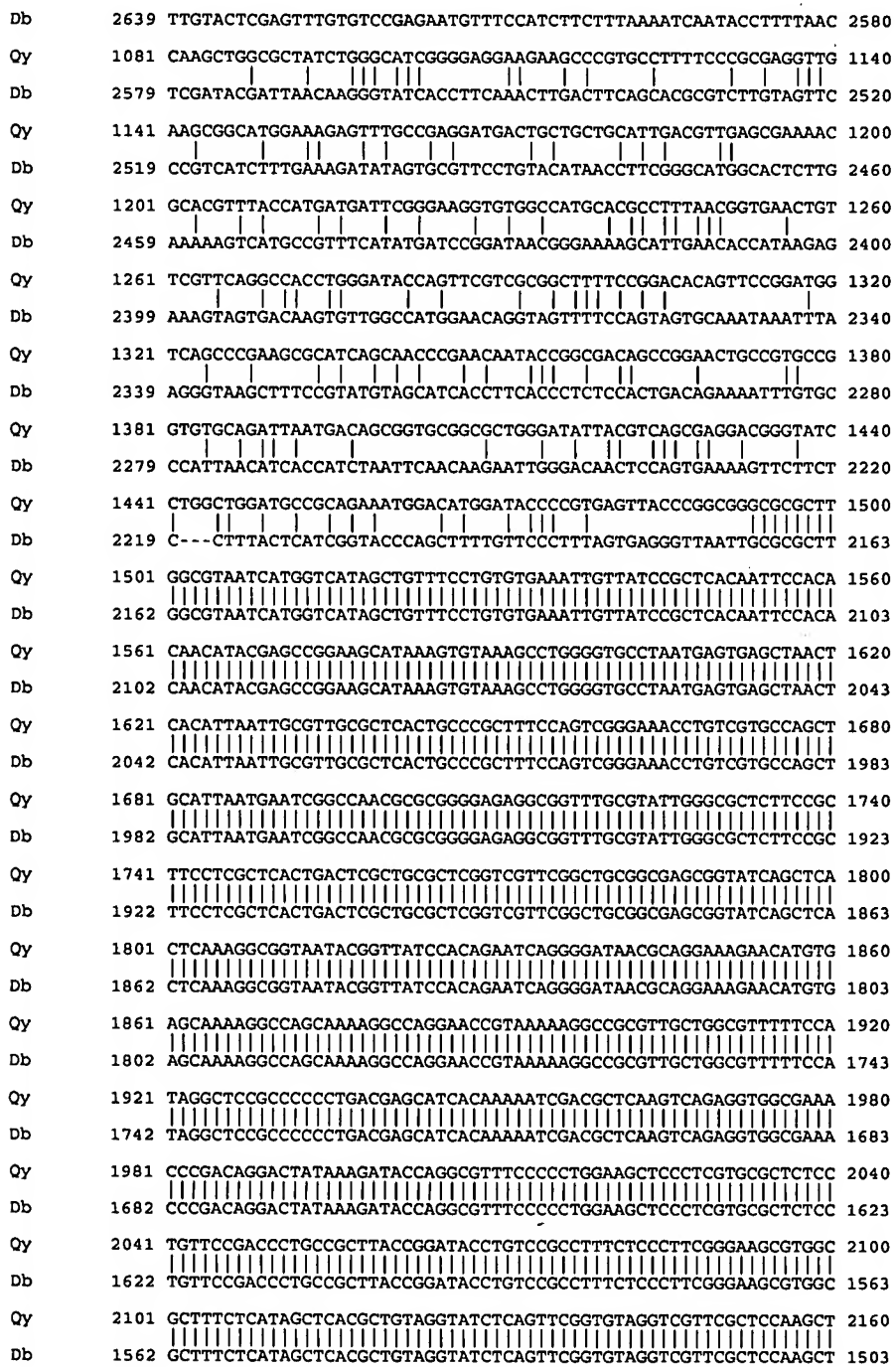
```
Qy       3480 GGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCT 3539
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 GGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCT 124

Qy       3540 TCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATAT 3599
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 TCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATAT 64

Qy       3600 TTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGC 3659
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          63 TTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGC 4

Qy       3660 CAC 3662
             |||
Db           3 CAC 1
```

RESULT 4
ABX14571/c
ID ABX14571 standard; DNA; 3637 BP.
XX
AC ABX14571;
XX
DT 14-MAR-2003 (first entry)
XX
DE Plasmid pBSC3GFP DNA.
XX
KW Circular; cyclic; heteroduplex; T4 DNA polymerase; T4 DNA ligase; CEL I;
KW population diversity; exonuclease; complementarity; ds.
XX
OS Synthetic.
XX
PN WO200279468-A2.
XX
PD 10-OCT-2002.
XX
PF 01-FEB-2002; 2002WO-US003055.
XX
PR 02-FEB-2001; 2001US-0266386P.
PR 14-FEB-2001; 2001US-0268785P.
XX
PA (LARG-) LARGE SCALE BIOLOGY CORP.
XX
PI Padgett HS, Fitzmaurice WP, Lindo JA;
XX
DR WPI; 2003-092898/08.
XX
PT Increasing complementarity by making sequence variants from heteroduplex
PT polynucleotides, useful for producing improved gene products from
PT randomly mutated genes.
XX
PS Example 5; Page 99-100; 105pp; English.
XX
CC The invention relates to an in vitro method of making sequence variants
CC from at least one heteroduplex polynucleotide that has at least two-non
CC complementary nucleotide base pairs. The method comprises combining a
CC heteroduplex polynucleotide with CEL I, T4 DNA polymerase and T4 DNA
CC ligase, or an agent or agents with exonuclease activity and allowing
CC sufficient time for the percentage of complementarity to increase, where
CC one or more variants are made. The method can also be used to increase
CC diversity in a population of sequences. The method is useful for
CC producing improved gene products from randomly mutated genes or from
CC expression from a suitable plant, animal, fungal, yeast or bacterial
CC expression vector, and provides a high-efficiency recovery of
CC recombinants for generating novel polynucleotides from parental templates
CC that are more diverse and with a lower percentage of sequence identity.
CC This sequence represents plasmid DNA, used in heteroduplex substrate
CC preparation
XX
SQ Sequence 3637 BP; 942 A; 892 C; 870 G; 933 T; 0 U; 0 Other;

Query Match 72.8%; Score 2664.8; DB 8; Length 3637;

Best Local Similarity 83.8%; Pred. No. 0;
Matches 3070; Conservative 0; Mismatches 567; Indels 25; Gaps 4;

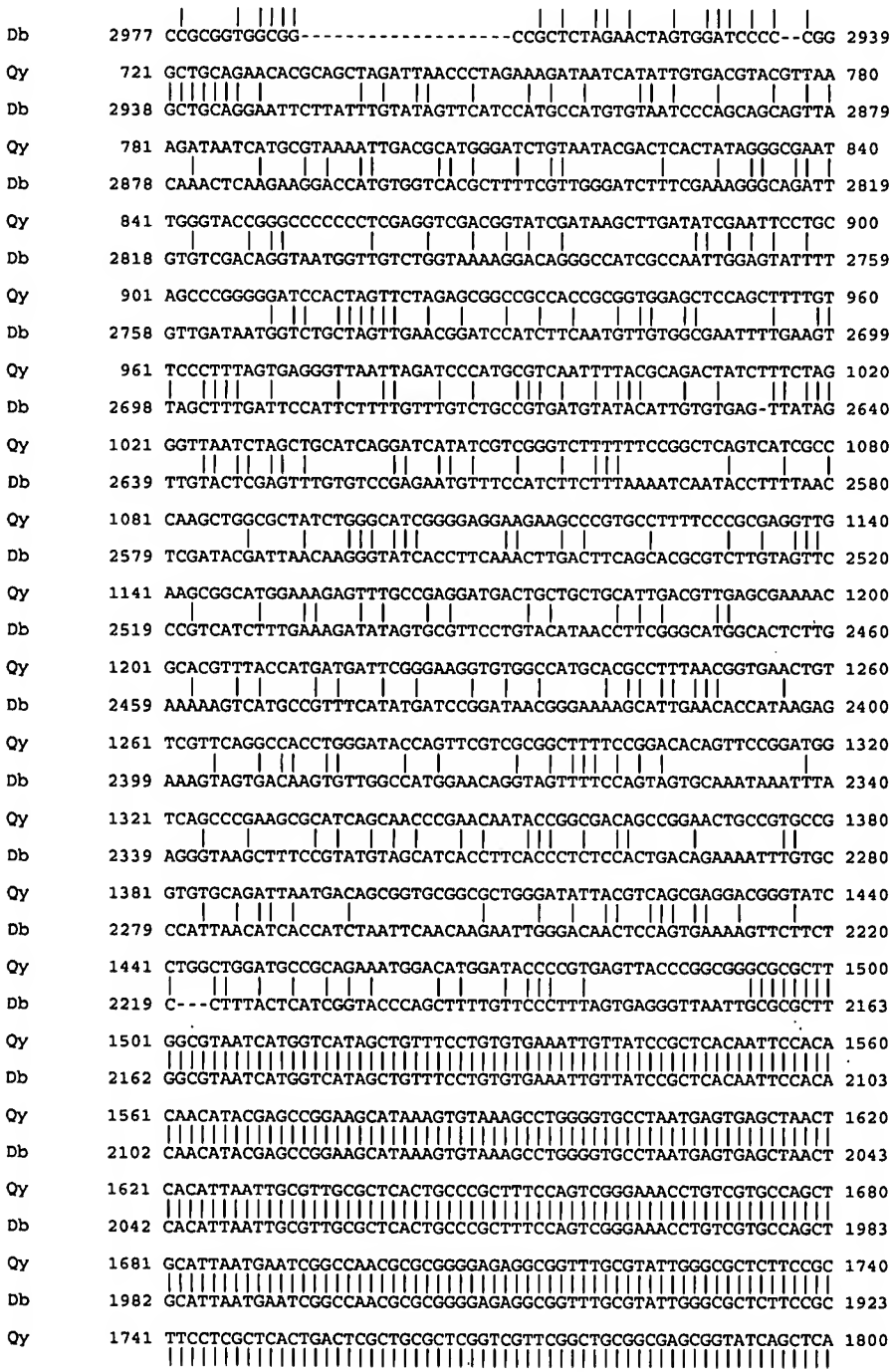
```
Qy          1 CTAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3637 CTAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTC 3578

Qy         61 ATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3577 ATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGA 3518

Qy        121 GATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3517 GATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTC 3458

Qy        181 CAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3457 CAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACC 3398

Qy        241 CTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3397 CTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAG 3338

Qy        301 CCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3337 CCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAA 3278

Qy        361 AGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCAC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3277 AGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCAC 3218

Qy        421 CACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCCCATTCGCCATTCAGGCTGCG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3217 CACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCCCATTCGCCATTCAGGCTGCG 3158

Qy        481 CAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3157 CAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGG 3098

Qy        541 GGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3097 GGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTG 3038

Qy        601 TAAAACGACGGCCAGTGAGCGCGCCTCGTTCATTCACGTTTTTGAACCCGTGGAGGACGG 660
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3037 TAAAACGACGGCCAGTGAGCGCGCGTAATACGACTCACTATAGGGCGAATTGGAGCTCCA 2978

Qy        661 GCAGACTCGCGGTGCAAATGTGTTTTACAGCGTGATGGAGCAGATGAAGATGCTCGACAC 720
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2977 CCGCGGTGGCGG----------CCGCTCTAGAACTAGTGGATCCCC--CGG 2939

Qy        721 GCTGCAGAACACGCAGCTAGATTAACCCTAGAAAGATAATCATATTGTGACGTACGTTAA 780
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2938 GCTGCAGGAATTCTTATTTGTATAGTTCATCCATGCCATGTGTAATCCCAGCAGCAGTTA 2879

Qy        781 AGATAATCATGCGTAAAATTGACGCATGGGATCTGTAATACGACTCACTATAGGGCGAAT 840
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2878 CAAACTCAAGAAGGACCATGTGGTCACGCTTTTCGTTGGGATCTTTCGAAAGGGCAGATT 2819

Qy        841 TGGGTACCGGGCCCCCCCCTCGAGGTCGACGGTATCGATAAGCTTGATATCGAATTCCTGC 900
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2818 GTGTCGACAGGTAATGGTTGTCTGGTAAAAGGACAGGGCCATCGCCAATTGGAGTATTTT 2759

Qy        901 AGCCCGGGGGATCCACTAGTTCTAGAGCGGCCGCCACCGCGGTGGAGCTCCAGCTTTTGT 960
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2758 GTTGATAATGGTCTGCTAGTTGAACGGATCCATCTTCAATGTTGTGGCGAATTTTGAAGT 2699

Qy        961 TCCCTTTAGTGAGGGTTAATTAGATCCCATGCGTCAATTTTACGCAGACTATCTTTCTAG 1020
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2698 TAGCTTTGATTCCATTCTTTTGTTTGTCTGCCGTGATGTATACATTGTGTGAG-TTATAG 2640

Qy       1021 GGTTAATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCC 1080
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```


Db 2639 TTGTACTCGAGTTTGTGTCCGAGAATGTTTCCATCTTCTTTAAAATCAATACCTTTTAAC 2580
Qy 1081 CAAGCTGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTG 1140
Db 2579 TCGATACGATTAACAAGGGTATCACCTTCAAACTTGACTTCAGCACGCGTCTTGTAGTTC 2520
Qy 1141 AAGCGGCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAAC 1200
Db 2519 CCGTCATCTTTGAAAGATATAGTGCGTTCCTGTACATAACCTTCGGGCATGGCACTCTTG 2460
Qy 1201 GCACGTTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGT 1260
Db 2459 AAAAAGTCATGCCGTTTCATATGATCCGGATAACGGGAAAAGCATTGAACACCATAAGAG 2400
Qy 1261 TCGTTCAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGG 1320
Db 2399 AAAGTAGTGACAAGTGTTGGCCATGGAACAGGTAGTTTTCCAGTAGTGCAAATAAATTTA 2340
Qy 1321 TCAGCCCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCG 1380
Db 2339 AGGGTAAGCTTTCCGTATGTAGCATCACCTTCACCCTCTCCACTGACAGAAAATTTGTGC 2280
Qy 1381 GTGTGCAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATC 1440
Db 2279 CCATTAACATCACCATCTAATTCAACAAGAATTGGGACAACTCCAGTGAAAAGTTCTTCT 2220
Qy 1441 CTGGCTGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCTT 1500
Db 2219 C---CTTTACTCATCGGTACCCAGCTTTTGTTCCCTTTAGTGAGGGTTAATTGCGCGCTT 2163
Qy 1501 GGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACA 1560
Db 2162 GGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACA 2103
Qy 1561 CAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACT 1620
Db 2102 CAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACT 2043
Qy 1621 CACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCT 1680
Db 2042 CACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCT 1983
Qy 1681 GCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGC 1740
Db 1982 GCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGC 1923
Qy 1741 TTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCA 1800
Db 1922 TTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCA 1863
Qy 1801 CTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTG 1860
Db 1862 CTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTG 1803
Qy 1861 AGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCA 1920
Db 1802 AGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCA 1743
Qy 1921 TAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAA 1980
Db 1742 TAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAA 1683
Qy 1981 CCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCC 2040
Db 1682 CCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCC 1623
Qy 2041 TGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGC 2100
Db 1622 TGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGC 1563
Qy 2101 GCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCT 2160
Db 1562 GCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCT 1503
Qy 2161 GGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCG 2220
Db 1502 GGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCG 1443
Qy 2221 TCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAG 2280
Db 1442 TCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAG 1383
Qy 2281 GATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTA 2340
Db 1382 GATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTA 1323
Qy 2341 CGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGG 2400
Db 1322 CGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGG 1263
Qy 2401 AAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTT 2460
Db 1262 AAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTT 1203
Qy 2461 TGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTT 2520
Db 1202 TGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTT 1143
Qy 2521 TTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAG 2580
Db 1142 TTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAG 1083
Qy 2581 ATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAAT 2640
Db 1082 ATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAAT 1023
Qy 2641 CTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACC 2700
Db 1022 CTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACC 963
Qy 2701 TATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGAT 2760
Db 962 TATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGAT 903
Qy 2761 AACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCC 2820
Db 902 AACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCC 843
Qy 2821 ACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAG 2880
Db 842 ACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAG 783
Qy 2881 AAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAG 2940
Db 782 AAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAG 723
Qy 2941 AGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGT 3000
Db 722 AGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGT 663
Qy 3001 GGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCG 3060
Db 662 GGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCG 603
Qy 3061 AGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGT 3120
Db 602 AGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGT 543
Qy 3121 TGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTC 3180
Db 542 TGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTC 483
Qy 3181 TCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTC 3240
Db 482 TCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTC 423

```
Qy       3241 ATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAA 3300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 ATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAA 363

Qy       3301 TACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCG 3360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 TACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCG 303

Qy       3361 AAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACC 3420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 AAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACC 243

Qy       3421 CAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAG 3480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 CAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAG 183

Qy       3481 GCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTT 3540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 GCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTT 123

Qy       3541 CCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATT 3600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 CCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATT 63

Qy       3601 TGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCC 3660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          62 TGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCC 3

Qy       3661 AC 3662
              ||
Db          2 AC 1
```

RESULT 5
ADM68448/c
ID ADM68448 standard; DNA; 3637 BP.
XX
AC ADM68448;
XX
DT 03-JUN-2004 (first entry)
XX
DE Cycle 3 green fluorescent protein, GFP, DNA construct.
XX
KW ds; mismatch endonuclease; endonuclease; gene shuffling technology;
KW single nucleotide polymorphism; cancer susceptibility;
KW sequence variation redistribution; cycle 3 GFP;
KW green fluorescent protein.
XX
OS Aequorea victoria.
OS Synthetic.
XX
PN US2003157682-A1.
XX
PD 21-AUG-2003.
XX
PF 31-JAN-2003; 2003US-00356708.
XX
PR 01-FEB-2002; 2002US-0353722P.
PR 14-MAR-2002; 2002US-00098155.
PR 01-AUG-2002; 2002US-00211079.
XX
PA (PADG/) PADGETT H S.
PA (VAEW/) VAEWHONGS A A.
PA (VOJD/) VOJDANI F S.
PA (SMIT/) SMITH M L.
PA (LIND/) LINDBO J A.
PA (FITZ/) FITZMAURICE W P.
XX
PI Padgett HS, Vaewhongs AA, Vojdani FS, Smith ML, Lindbo JA;
PI Fitzmaurice WP;
XX
DR WPI; 2003-766176/72.
XX
PT Making a mismatch endonuclease, useful in gene shuffling and in detection
PT of single nucleotide polymorphisms, comprises transfecting a host with a
PT recombinant viral vector including a polynucleotide encoding a mismatch
PT endonuclease.
XX
PS Example 9; SEQ ID NO 17; 79pp; English.
XX
CC The invention relates to a method of making a mismatch endonuclease
CC enzyme comprising transfecting a host plant, animal, yeast, fungus or
CC bacterium with a recombinant viral vector that encodes a polynucleotide
CC sequence for a mismatch endonuclease, growing the host so that the
CC polynucleotide is expressed, and extracting the mismatch endonuclease
CC enzyme from the host. The method is useful for making mismatch
CC endonuclease enzymes, for obtaining peptides and polynucleotides with
CC desired functional properties and for detecting mutations. The mismatch
CC endonuclease enzymes are useful in gene shuffling technology for
CC developing new genes, in detecting single nucleotide polymorphisms for
CC e.g. detecting evidence of cancer susceptibility, or in redistributing
CC sequence variations between non-identical polynucleotide sequences. The
CC present sequence represents the cycle 3 green fluorescent protein, GFP,
CC DNA construct.
XX
SQ Sequence 3637 BP; 942 A; 892 C; 870 G; 933 T; 0 U; 0 Other;

```
  Query Match             72.8%;  Score 2664.8;  DB 11;  Length 3637;
  Best Local Similarity   83.8%;  Pred. No. 0;
  Matches 3070;  Conservative    0;  Mismatches  567;  Indels   25;  Gaps    4;

Qy          1 CTAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3637 CTAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTC 3578

Qy         61 ATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3577 ATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGA 3518

Qy        121 GATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3517 GATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTC 3458

Qy        181 CAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3457 CAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACC 3398

Qy        241 CTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3397 CTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAG 3338

Qy        301 CCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3337 CCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAA 3278

Qy        361 AGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCAC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3277 AGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCAC 3218

Qy        421 CACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCCCATTCGCCATTCAGGCTGCG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3217 CACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCCCATTCGCCATTCAGGCTGCG 3158

Qy        481 CAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3157 CAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGG 3098

Qy        541 GGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3097 GGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTG 3038

Qy        601 TAAAACGACGGCCAGTGAGCGCGCCTCGTTCATTCACGTTTTTGAACCCGTGGAGGACGG 660
              |||||||||||||||||||||||| | | |  | | |  |||||  |
Db       3037 TAAAACGACGGCCAGTGAGCGCGCGTAATACGACTCACTATAGGGCGAATTGGAGCTCCA 2978

Qy        661 GCAGACTCGCGGTGCAAATGTGTTTTACAGCGTGATGGAGCAGATGAAGATGCTCGACAC 720
```

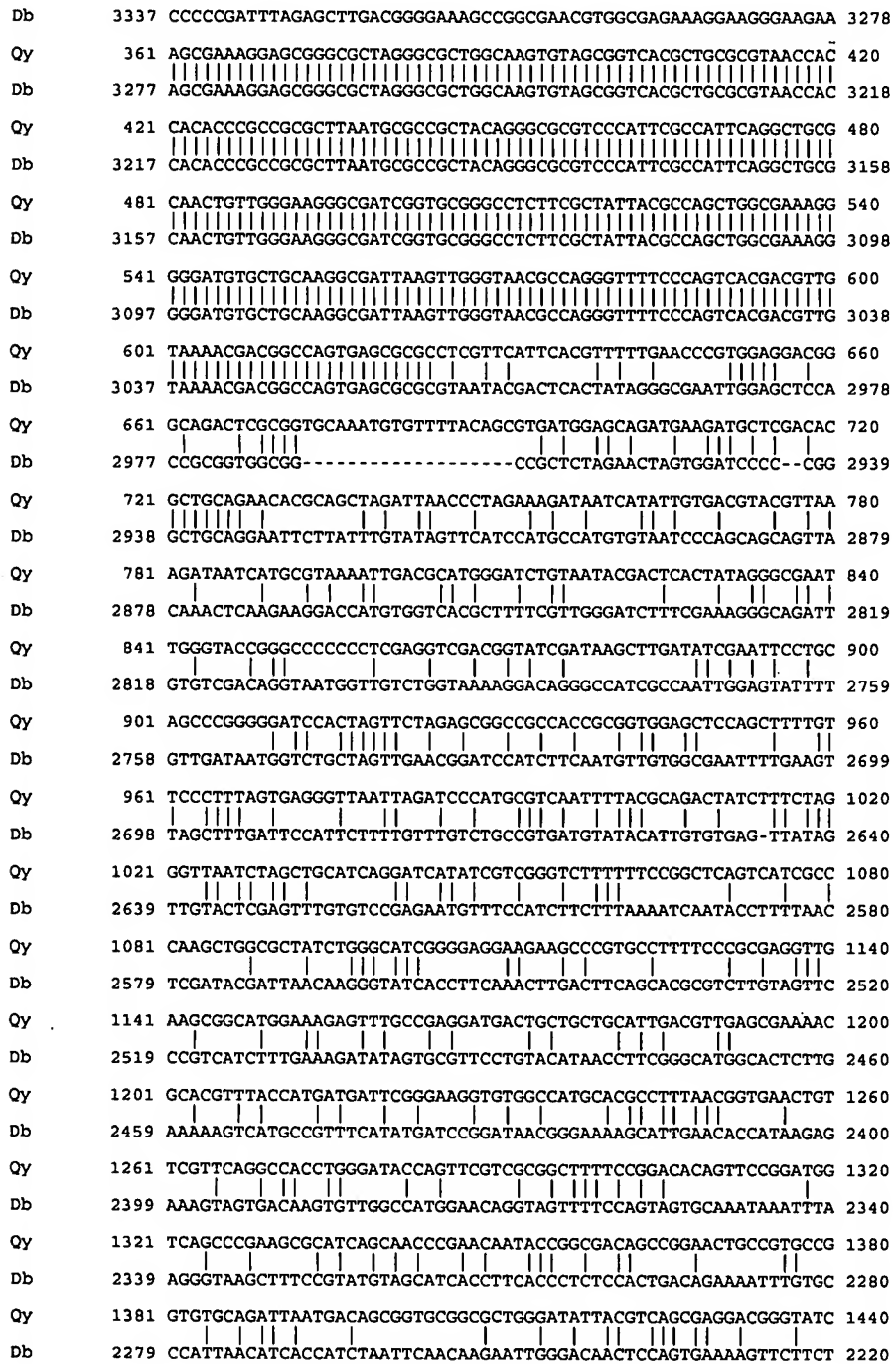

Db 2977 CCGCGGTGGCGG----------CCGCTCTAGAACTAGTGGATCCCC--CGG 2939
Qy 721 GCTGCAGAACACGCAGCTAGATTAACCCTAGAAAGATAATCATATTGTGACGTACGTTAA 780
Db 2938 GCTGCAGGAATTCTTATTTGTATAGTTCATCCATGCCATGTGTAATCCCAGCAGCAGTTA 2879
Qy 781 AGATAATCATGCGTAAAATTGACGCATGGGATCTGTAATACGACTCACTATAGGGCGAAT 840
Db 2878 CAAACTCAAGAAGGACCATGTGGTCACGCTTTTCGTTGGGATCTTTCGAAAGGGCAGATT 2819
Qy 841 TGGGTACCGGGCCCCCCCTCGAGGTCGACGGTATCGATAAGCTTGATATCGAATTCCTGC 900
Db 2818 GTGTCGACAGGTAATGGTTGTCTGGTAAAAGGACAGGGCCATCGCCAATTGGAGTATTTT 2759
Qy 901 AGCCCGGGGGATCCACTAGTTCTAGAGCGGCCGCCACCGCGGTGGAGCTCCAGCTTTTGT 960
Db 2758 GTTGATAATGGTCTGCTAGTTGAACGGATCCATCTTCAATGTTGTGGCGAATTTTGAAGT 2699
Qy 961 TCCCTTTAGTGAGGGTTAATTAGATCCCATGCGTCAATTTTACGCAGACTATCTTTCTAG 1020
Db 2698 TAGCTTTGATTCCATTCTTTTGTTTGTCTGCCGTGATGTATACATTGTGTGAG-TTATAG 2640
Qy 1021 GGTTAATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCC 1080
Db 2639 TTGTACTCGAGTTTGTGTCCGAGAATGTTTCCATCTTCTTTAAAATCAATACCTTTTAAC 2580
Qy 1081 CAAGCTGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTG 1140
Db 2579 TCGATACGATTAACAAGGGTATCACCTTCAAACTTGACTTCAGCACGCGTCTTGTAGTTC 2520
Qy 1141 AAGCGGCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAAC 1200
Db 2519 CCGTCATCTTTGAAAGATATAGTGCGTTCCTGTACATAACCTTCGGGCATGGCACTCTTG 2460
Qy 1201 GCACGTTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGT 1260
Db 2459 AAAAAGTCATGCCGTTTCATATGATCCGGATAACGGGAAAAGCATTGAACACCATAAGAG 2400
Qy 1261 TCGTTCAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGG 1320
Db 2399 AAAGTAGTGACAAGTGTTGGCCATGGAACAGGTAGTTTTCCAGTAGTGCAAATAAATTTA 2340
Qy 1321 TCAGCCCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCG 1380
Db 2339 AGGGTAAGCTTTCCGTATGTAGCATCACCTTCACCCTCTCCACTGACAGAAAATTTGTGC 2280
Qy 1381 GTGTGCAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATC 1440
Db 2279 CCATTAACATCACCATCTAATTCAACAAGAATTGGGACAACTCCAGTGAAAAGTTCTTCT 2220
Qy 1441 CTGGCTGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCTT 1500
Db 2219 C---CTTTACTCATCGGTACCCAGCTTTTGTTCCCTTTAGTGAGGGTTAATTGCGCGCTT 2163
Qy 1501 GGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACA 1560
Db 2162 GGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACA 2103
Qy 1561 CAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACT 1620
Db 2102 CAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACT 2043
Qy 1621 CACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCT 1680
Db 2042 CACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCT 1983
Qy 1681 GCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGC 1740
Db 1982 GCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGC 1923
Qy 1741 TTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCA 1800
Db 1922 TTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCA 1863
Qy 1801 CTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTG 1860
Db 1862 CTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTG 1803
Qy 1861 AGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCA 1920
Db 1802 AGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCA 1743
Qy 1921 TAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAA 1980
Db 1742 TAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAA 1683
Qy 1981 CCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCC 2040
Db 1682 CCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCC 1623
Qy 2041 TGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGC 2100
Db 1622 TGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGC 1563
Qy 2101 GCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCT 2160
Db 1562 GCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCT 1503
Qy 2161 GGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCG 2220
Db 1502 GGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCG 1443
Qy 2221 TCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAG 2280
Db 1442 TCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAG 1383
Qy 2281 GATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTA 2340
Db 1382 GATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTA 1323
Qy 2341 CGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGG 2400
Db 1322 CGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGG 1263
Qy 2401 AAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTT 2460
Db 1262 AAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTT 1203
Qy 2461 TGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTT 2520
Db 1202 TGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTT 1143
Qy 2521 TTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAG 2580
Db 1142 TTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAG 1083
Qy 2581 ATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAAT 2640
Db 1082 ATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAAT 1023
Qy 2641 CTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACC 2700
Db 1022 CTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACC 963
Qy 2701 TATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGAT 2760
Db 962 TATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGAT 903
Qy 2761 AACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCC 2820
Db 902 AACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCC 843
Qy 2821 ACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAG 2880
Db 842 ACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAG 783

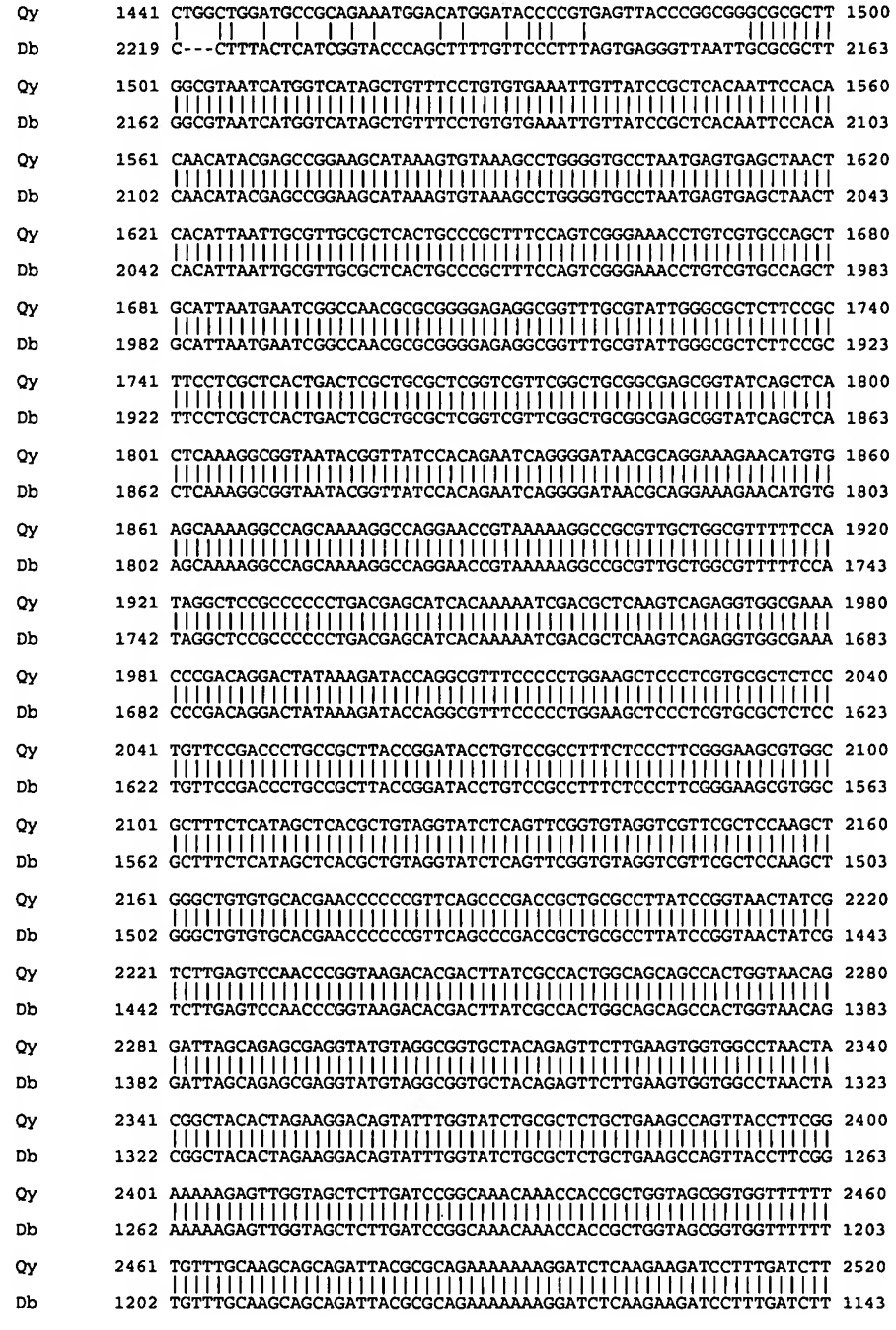
```
Qy       2881 AAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAG 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        782 AAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAG 723

Qy       2941 AGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGT 3000
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        722 AGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGT 663

Qy       3001 GGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCG 3060
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        662 GGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCG 603

Qy       3061 AGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGT 3120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        602 AGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGT 543

Qy       3121 TGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTC 3180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        542 TGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTC 483

Qy       3181 TCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTC 3240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        482 TCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTC 423

Qy       3241 ATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAA 3300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 ATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAA 363

Qy       3301 TACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCG 3360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 TACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCG 303

Qy       3361 AAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACC 3420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 AAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACC 243

Qy       3421 CAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAG 3480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 CAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAG 183

Qy       3481 GCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTT 3540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 GCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTT 123

Qy       3541 CCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATT 3600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 CCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATT 63

Qy       3601 TGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCC 3660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          62 TGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCC 3

Qy       3661 AC 3662
              ||
Db          2 AC 1
```

RESULT 6
ADP26601/c
ID ADP26601 standard; DNA; 3637 BP.
XX
AC ADP26601;
XX
DT 26-AUG-2004 (first entry)
XX
DE Green fluorescent protein (GFP) plasmid DNA #2.
XX
KW Sequence variation; heteroduplex; transcription; DNA integration;
KW ribozyme expression; gene; ds; green fluorescent protein; GFP.
XX
OS Aequorea victoria.
OS Synthetic.
XX
PN US2004110130-A1.
XX
PD 10-JUN-2004.
XX
PF 25-OCT-2002; 2002US-00280913.
XX
PR 02-FEB-2001; 2001US-0266386P.
PR 14-FEB-2001; 2001US-0268785P.
PR 01-FEB-2002; 2002US-00066390.
PR 08-AUG-2002; 2002US-0402342P.
XX
PA (LARG-) LARGE SCALE BIOLOGY CORP.
XX
PI Padgett HS, Lindbo JA, Fitzmaurice WP;
XX
DR WPI; 2004-440326/41.
XX
PT Redistributing sequence variations between non-identical polynucleotide
PT sequences, useful for generating improved polynucleotide having a desired
PT characteristic, comprises making a heteroduplex and introducing a nick.
XX
PS Example 5; SEQ ID NO 17; 75pp; English.
XX
CC The invention relates to an in vitro method of redistributing sequence
CC variations between non-identical polynucleotide sequences, comprising
CC making a heteroduplex polynucleotide from two non-identical
CC polynucleotides, introducing a nick in the second strand at or near a
CC base pair mismatch site, removing the mismatched base(s) from the
CC mismatch site where the nick occurred and using the first strand as a
CC template to replace the removed base(s) with bases that complement the
CC base(s) in the first strand. The invention also relates to an in vitro
CC method of making a population of sequence variants from a heteroduplex
CC polynucleotide sequence, obtaining a polynucleotide sequence encoding a
CC desired functional property and identifying a reasserted DNA molecule
CC encoding a protein with a desired functional property. The method is
CC useful for generating an improved polynucleotide sequence or a population
CC of improved polynucleotide sequences possessing at least one desired
CC phenotypic characteristic (e.g., promotes transcription of linked
CC polynucleotides), where such polynucleotides are useful for expression
CC from a plant, animal, fungal, yeast, or bacterial expression vector, for
CC integration to form a transgenic plant, animal or microorganism, and for
CC expression of a ribozyme. This sequence represents DNA used in the scope
CC of the invention.
XX
SQ Sequence 3637 BP; 942 A; 892 C; 870 G; 933 T; 0 U; 0 Other;

Query Match   72.8%; Score 2664.8; DB 12; Length 3637;
Best Local Similarity   83.8%; Pred. No. 0;
Matches 3070; Conservative   0; Mismatches 567; Indels   25; Gaps   4;

```
Qy          1 CTAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3637 CTAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTC 3578

Qy         61 ATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3577 ATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGA 3518

Qy        121 GATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3517 GATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTC 3458

Qy        181 CAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3457 CAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACC 3398

Qy        241 CTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3397 CTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAG 3338

Qy        301 CCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```



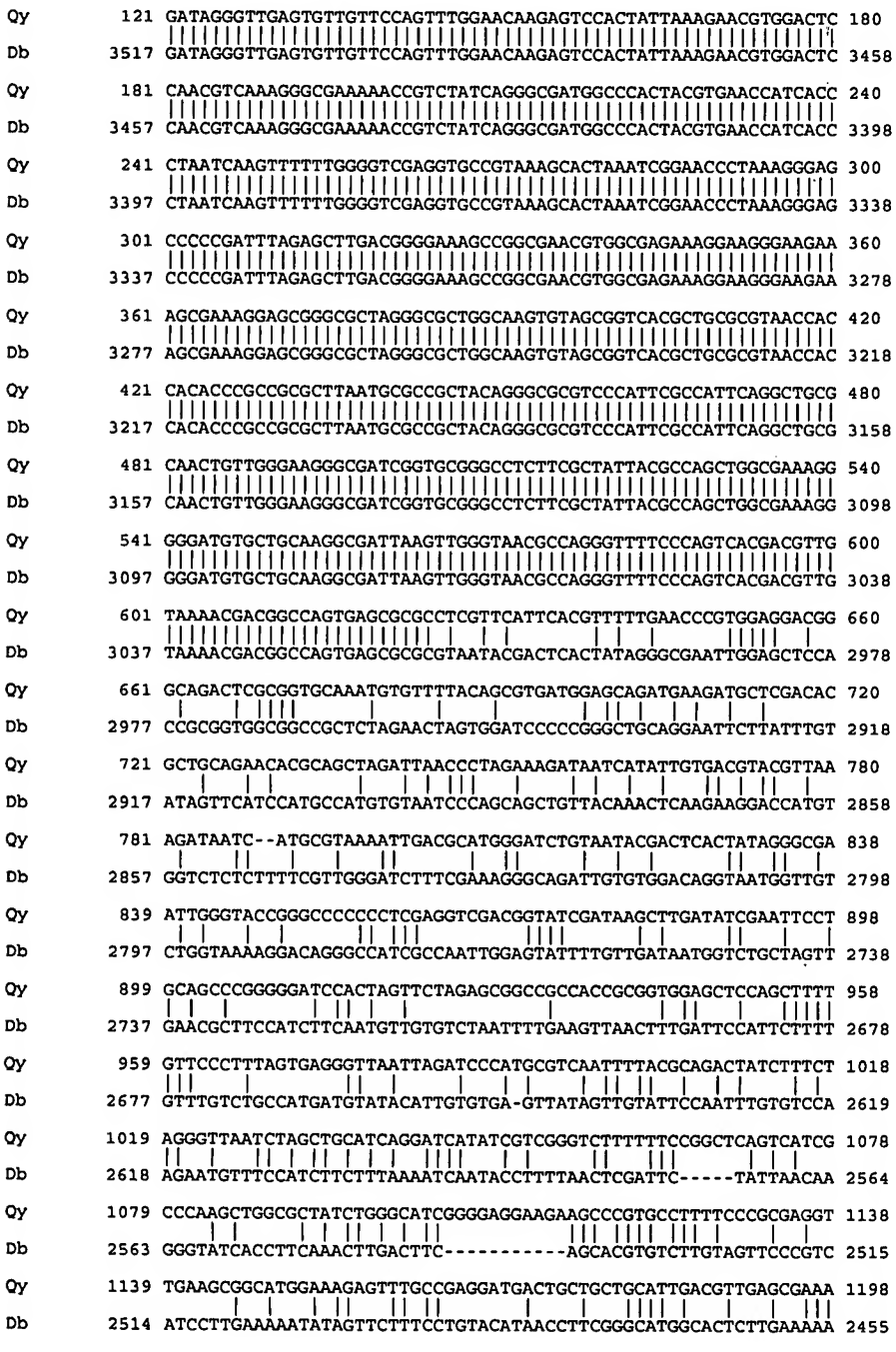


```
Qy       2521 TTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAG 2580
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1142 TTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAG 1083

Qy       2581 ATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAAT 2640
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1082 ATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAAT 1023

Qy       2641 CTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACC 2700
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1022 CTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACC 963

Qy       2701 TATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGAT 2760
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        962 TATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGAT 903

Qy       2761 AACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCC 2820
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        902 AACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCC 843

Qy       2821 ACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAG 2880
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        842 ACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAG 783

Qy       2881 AAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAG 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        782 AAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAG 723

Qy       2941 AGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGT 3000
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        722 AGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGT 663

Qy       3001 GGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCG 3060
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        662 GGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCG 603

Qy       3061 AGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGT 3120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        602 AGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGT 543

Qy       3121 TGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTC 3180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        542 TGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTC 483

Qy       3181 TCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTC 3240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        482 TCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTC 423

Qy       3241 ATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAA 3300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 ATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAA 363

Qy       3301 TACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCG 3360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 TACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCG 303

Qy       3361 AAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACC 3420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 AAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACC 243

Qy       3421 CAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAG 3480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 CAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAG 183

Qy       3481 GCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTT 3540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 GCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTT 123

Qy       3541 CCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATT 3600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 CCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATT 63

Qy       3601 TGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCC 3660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          62 TGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCC 3

Qy       3661 AC 3662
              ||
Db          2 AC 1
```

RESULT 7
ABX14570/c
ID ABX14570 standard; DNA; 3637 BP.
XX
AC ABX14570;
XX
DT 14-MAR-2003 (first entry)
XX
DE Plasmid pBSWTGFP DNA.
XX
KW Circular; cyclic; heteroduplex; T4 DNA polymerase; T4 DNA ligase; CEL I;
KW population diversity; exonuclease; complementarity; ds.
XX
OS Synthetic.
XX
PN WO200279468-A2.
XX
PD 10-OCT-2002.
XX
PF 01-FEB-2002; 2002WO-US003055.
XX
PR 02-FEB-2001; 2001US-0266386P.
PR 14-FEB-2001; 2001US-0268785P.
XX
PA (LARG-) LARGE SCALE BIOLOGY CORP.
XX
PI Padgett HS, Fitzmaurice WP, Lindo JA;
XX
DR WPI; 2003-092898/08.
XX
PT Increasing complementarity by making sequence variants from heteroduplex
PT polynucleotides, useful for producing improved gene products from
PT randomly mutated genes.
XX
PS Example 5; Page 98-99; 105pp; English.
XX
CC The invention relates to an in vitro method of making sequence variants
CC from at least one heteroduplex polynucleotide that has at least two-non
CC complementary nucleotide base pairs. The method comprises combining a
CC heteroduplex polynucleotide with CEL I, T4 DNA polymerase and T4 DNA
CC ligase, or an agent or agents with exonuclease activity and allowing
CC sufficient time for the percentage of complementarity to increase, where
CC one or more variants are made. The method can also be used to increase
CC diversity in a population of sequences. The method is useful for
CC producing improved gene products from randomly mutated genes or from
CC expression from a suitable plant, animal, fungal, yeast or bacterial
CC expression vector, and provides a high-efficiency recovery of
CC recombinants for generating novel polynucleotides from parental templates
CC that are more diverse and with a lower percentage of sequence identity.
CC This sequence represents plasmid DNA, used in heteroduplex substrate
CC preparation
XX
SQ Sequence 3637 BP; 959 A; 879 C; 868 G; 931 T; 0 U; 0 Other;

Query Match 72.5%; Score 2656.4; DB 8; Length 3637;
Best Local Similarity 84.0%; Pred. No. 0;
Matches 3079; Conservative 0; Mismatches 556; Indels 29; Gaps 6;

```
Qy          1 CTAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3637 CTAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTC 3578

Qy         61 ATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3577 ATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGA 3518
```


Qy 121 GATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTC 180
Db 3517 GATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTC 3458
Qy 181 CAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACC 240
Db 3457 CAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACC 3398
Qy 241 CTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAG 300
Db 3397 CTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAG 3338
Qy 301 CCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAA 360
Db 3337 CCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAA 3278
Qy 361 AGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCAC 420
Db 3277 AGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCAC 3218
Qy 421 CACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCCCATTCGCCATTCAGGCTGCG 480
Db 3217 CACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCCCATTCGCCATTCAGGCTGCG 3158
Qy 481 CAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGG 540
Db 3157 CAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGG 3098
Qy 541 GGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTG 600
Db 3097 GGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTG 3038
Qy 601 TAAAACGACGGCCAGTGAGCGCGCCTCGTTCATTCACGTTTTTGAACCCGTGGAGGACGG 660
Db 3037 TAAAACGACGGCCAGTGAGCGCGCGTAATACGACTCACTATAGGGCGAATTGGAGCTCCA 2978
Qy 661 GCAGACTCGCGGTGCAAATGTGTTTTACAGCGTGATGGAGCAGATGAAGATGCTCGACAC 720
Db 2977 CCGCGGTGGCGGCCGCTCTAGAACTAGTGGATCCCCCGGGCTGCAGGAATTCTTATTTGT 2918
Qy 721 GCTGCAGAACACGCAGCTAGATTAACCCTAGAAAGATAATCATATTGTGACGTACGTTAA 780
Db 2917 ATAGTTCATCCATGCCATGTGTAATCCCAGCAGCTGTTACAAACTCAAGAAGGACCATGT 2858
Qy 781 AGATAATC--ATGCGTAAAATTGACGCATGGGATCTGTAATACGACTCACTATAGGGCGA 838
Db 2857 GGTCTCTCTTTTCGTTGGGATCTTTCGAAAGGGCAGATTGTGTGGACAGGTAATGGTTGT 2798
Qy 839 ATTGGGTACCGGGCCCCCCCTCGAGGTCGACGGTATCGATAAGCTTGATATCGAATTCCT 898
Db 2797 CTGGTAAAAGGACAGGGCCATCGCCAATTGGAGTATTTTGTTGATAATGGTCTGCTAGTT 2738
Qy 899 GCAGCCCGGGGGATCCACTAGTTCTAGAGCGGCCGCCACCGCGGTGGAGCTCCAGCTTTT 958
Db 2737 GAACGCTTCCATCTTCAATGTTGTGTCTAATTTTGAAGTTAACTTTGATTCCATTCTTTT 2678
Qy 959 GTTCCCTTTAGTGAGGGTTAATTAGATCCCATGCGTCAATTTTACGCAGACTATCTTTCT 1018
Db 2677 GTTTGTCTGCCATGATGTATACATTGTGTGA-GTTATAGTTGTATTCCAATTTGTGTCCA 2619
Qy 1019 AGGGTTAATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCG 1078
Db 2618 AGAATGTTTCCATCTTCTTTAAAATCAATACCTTTTAACTCGATTC-----TATTAACAA 2564
Qy 1079 CCCAAGCTGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGT 1138
Db 2563 GGGTATCACCTTCAAACTTGACTTC----------AGCACGTGTCTTGTAGTTCCCGTC 2515
Qy 1139 TGAAGCGGCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAA 1198
Db 2514 ATCCTTGAAAAATATAGTTCTTTCCTGTACATAACCTTCGGGCATGGCACTCTTGAAAAA 2455
Qy 1199 ACGCACGTTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACT 1258
Db 2454 GTCATGCCGTTTCATATGATCTGGGTATCTTGAAAAGCATTGAACACCATAAGAGAAAGT 2395
Qy 1259 GTTCGTTCAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGAT 1318
Db 2394 AGTGACAAGTGTTGGCCATGGAACAGGTAGTTTTCCAGTAGTGCAAATAAATTTAAGGGT 2335
Qy 1319 GGTCAGCCCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGC 1378
Db 2334 AAGTTTTC-----CGTATGTTGCATCACCTTCACCCTCTCCACTGACAGAAAATTTGTGC 2280
Qy 1379 CGGTGTGCAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTA 1438
Db 2279 CCATTAACATCACCATCTAATTCAACAAGAATTGGGACAACTCCAGTGAAAAGTTCTTCT 2220
Qy 1439 TCCTGGCTGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGC 1498
Db 2219 CCTTTACTCATCGGTACCCAGCTTTGTTCCCTTTAGTGAGGGTTA-----ATTGCGCGC 2165
Qy 1499 TTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCA 1558
Db 2164 TTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCA 2105
Qy 1559 CACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAA 1618
Db 2104 CACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAA 2045
Qy 1619 CTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAG 1678
Db 2044 CTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAG 1985
Qy 1679 CTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCC 1738
Db 1984 CTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCC 1925
Qy 1739 GCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCT 1798
Db 1924 GCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCT 1865
Qy 1799 CACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATG 1858
Db 1864 CACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATG 1805
Qy 1859 TGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTC 1918
Db 1804 TGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTC 1745
Qy 1919 CATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGA 1978
Db 1744 CATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGA 1685
Qy 1979 AACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCT 2038
Db 1684 AACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCT 1625
Qy 2039 CCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTG 2098
Db 1624 CCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTG 1565
Qy 2099 GCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAG 2158
Db 1564 GCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAG 1505
Qy 2159 CTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTAT 2218
Db 1504 CTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTAT 1445
Qy 2219 CGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAAC 2278
Db 1444 CGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAAC 1385
Qy 2279 AGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAAC 2338

```
Db       1384 AGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAAC 1325
Qy       2339 TACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTC 2398
Db       1324 TACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTC 1265
Qy       2399 GGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTT 2458
Db       1264 GGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTT 1205
Qy       2459 TTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATC 2518
Db       1204 TTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATC 1145
Qy       2519 TTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATG 2578
Db       1144 TTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATG 1085
Qy       2579 AGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCA 2638
Db       1084 AGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCA 1025
Qy       2639 ATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCA 2698
Db       1024 ATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCA 965
Qy       2699 CCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAG 2758
Db        964 CCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAG 905
Qy       2759 ATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGAC 2818
Db        904 ATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGAC 845
Qy       2819 CCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGC 2878
Db        844 CCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGC 785
Qy       2879 AGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCT 2938
Db        784 AGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCT 725
Qy       2939 AGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATC 2998
Db        724 AGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATC 665
Qy       2999 GTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGG 3058
Db        664 GTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGG 605
Qy       3059 CGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATC 3118
Db        604 CGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATC 545
Qy       3119 GTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAAT 3178
Db        544 GTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAAT 485
Qy       3179 TCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAG 3238
Db        484 TCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAG 425
Qy       3239 TCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGAT 3298
Db        424 TCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGAT 365
Qy       3299 AATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGG 3358
Db        364 AATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGG 305
Qy       3359 CGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCA 3418
Db        304 CGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCA 245
Qy       3419 CCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGA 3478
Db        244 CCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGA 185
Qy       3479 AGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTC 3538
Db        184 AGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTC 125
Qy       3539 TTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATA 3598
Db        124 TTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATA 65
Qy       3599 TTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTG 3658
Db          64 TTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTG 5
Qy       3659 CCAC 3662
Db           4 CCAC 1
```

RESULT 8
ADM68462/c
ID ADM68462 standard; DNA; 3637 BP.
XX
AC ADM68462;
XX
DT 03-JUN-2004 (first entry)
XX
DE Wild type green fluorescent protein, GFP, gene construct.
XX
KW ds; mismatch endonuclease; endonuclease; gene shuffling technology;
KW single nucleotide polymorphism; cancer susceptibility;
KW sequence variation redistribution; GFP; green fluorescent protein; gene.
XX
OS Aequorea victoria.
OS Synthetic.
XX
PN US2003157682-A1.
XX
PD 21-AUG-2003.
XX
PF 31-JAN-2003; 2003US-00356708.
XX
PR 01-FEB-2002; 2002US-0353722P.
PR 14-MAR-2002; 2002US-00098155.
PR 01-AUG-2002; 2002US-00211079.
XX
PA (PADG/) PADGETT H S.
PA (VAEW/) VAEWHONGS A A.
PA (VOJD/) VOJDANI F S.
PA (SMIT/) SMITH M L.
PA (LIND/) LINDBO J A.
PA (FITZ/) FITZMAURICE W P.
XX
PI Padgett HS, Vaewhongs AA, Vojdani FS, Smith ML, Lindbo JA;
PI Fitzmaurice WP;
XX
DR WPI; 2003-766176/72.
XX
PT Making a mismatch endonuclease, useful in gene shuffling and in detection
PT of single nucleotide polymorphisms, comprises transfecting a host with a
PT recombinant viral vector including a polynucleotide encoding a mismatch
PT endonuclease.
XX
PS Example 5; SEQ ID NO 31; 79pp; English.
XX
CC The invention relates to a method of making a mismatch endonuclease
CC enzyme comprising transfecting a host plant, animal, yeast, fungus or
CC bacterium with a recombinant viral vector that encodes a polynucleotide
CC sequence for a mismatch endonuclease, growing the host so that the

```
CC   polynucleotide is expressed, and extracting the mismatch endonuclease
CC   enzyme from the host. The method is useful for making mismatch
CC   endonuclease enzymes, for obtaining peptides and polynucleotides with
CC   desired functional properties and for detecting mutations. The mismatch
CC   endonuclease enzymes are useful in gene shuffling technology for
CC   developing new genes, in detecting single nucleotide polymorphisms for
CC   e.g. detecting evidence of cancer susceptibility, or in redistributing
CC   sequence variations between non-identical polynucleotide sequences. The
CC   present sequence represents the wild type green fluorescent protein, GFP,
CC   gene construct.
XX
SQ   Sequence 3637 BP; 959 A; 879 C; 868 G; 931 T; 0 U; 0 Other;

  Query Match            72.5%;  Score 2656.4;  DB 11;  Length 3637;
  Best Local Similarity  84.0%;  Pred. No. 0;
  Matches 3079;  Conservative    0;  Mismatches  556;  Indels   29;  Gaps    6;

Qy          1 CTAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3637 CTAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTC 3578

Qy         61 ATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3577 ATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGA 3518

Qy        121 GATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3517 GATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTC 3458

Qy        181 CAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3457 CAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACC 3398

Qy        241 CTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3397 CTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAG 3338

Qy        301 CCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3337 CCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAA 3278

Qy        361 AGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCAC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3277 AGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCAC 3218

Qy        421 CACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCCCATTCGCCATTCAGGCTGCG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3217 CACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCCCATTCGCCATTCAGGCTGCG 3158

Qy        481 CAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3157 CAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGG 3098

Qy        541 GGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3097 GGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTG 3038

Qy        601 TAAAACGACGGCCAGTGAGCGCGCCTCGTTCATTCACGTTTTTGAACCCGTGGAGGACGG 660
              |||||||||||||||||||||| | | |  | | |    |||||  |
Db       3037 TAAAACGACGGCCAGTGAGCGCGCGTAATACGACTCACTATAGGGCGAATTGGAGCTCCA 2978

Qy        661 GCAGACTCGCGGTGCAAATGTGTTTTACAGCGTGATGGAGCAGATGAAGATGCTCGACAC 720
              | |  |||| |  |  |  |  | || | | | |  |
Db       2977 CCGCGGTGGCGGCCGCTCTAGAACTAGTGGATCCCCCGGGCTGCAGGAATTCTTATTTGT 2918

Qy        721 GCTGCAGAACACGCAGCTAGATTAACCCTAGAAAGATAATCATATTGTGACGTACGTTAA 780
              |  | |         |  | | |||  |  | |  |  | || | |  |
Db       2917 ATAGTTCATCCATGCCATGTGTAATCCCAGCAGCTGTTACAAACTCAAGAAGGACCATGT 2858

Qy        781 AGATAATC--ATGCGTAAAATTGACGCATGGGATCTGTAATACGACTCACTATAGGGCGA 838
              |  ||  |  |  ||      | ||    | | |  |  || ||  |
Db       2857 GGTCTCTCTTTTCGTTGGGATCTTTCGAAAGGGCAGATTGTGTGGACAGGTAATGGTTGT 2798

Qy        839 ATTGGGTACCGGGCCCCCCCTCGAGGTCGACGGTATCGATAAGCTTGATATCGAATTCCT 898
              | | | |  |  || |||      ||||     | |  ||  |  |
Db       2797 CTGGTAAAAGGACAGGGCCATCGCCAATTGGAGTATTTTGTTGATAATGGTCTGCTAGTT 2738

Qy        899 GCAGCCCGGGGGATCCACTAGTTCTAGAGCGGCCGCCACCGCGGTGGAGCTCCAGCTTTT 958
              | | |       | || | |  |                 |  ||  |  |||||
Db       2737 GAACGCTTCCATCTTCAATGTTGTGTCTAATTTTGAAGTTAACTTTGATTCCATTCTTTT 2678

Qy        959 GTTCCCTTTAGTGAGGGTTAATTAGATCCCATGCGTCAATTTTACGCAGACTATCTTTCT 1018
              |||    |     ||  |    | | |   | || ||  | | |   | |
Db       2677 GTTTGTCTGCCATGATGTATACATTGTGTGA-GTTATAGTTGTATTCCAATTTGTGTCCA 2619

Qy       1019 AGGGTTAATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCG 1078
              ||  |  || | || | | |  ||||  | |    ||  |||     | | |
Db       2618 AGAATGTTTCCATCTTCTTTAAAATCAATACCTTTTAACTCGATTC-----TATTAACAA 2564

Qy       1079 CCCAAGCTGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGT 1138
                | |     | | || | | ||          ||| |||| ||| |    | |
Db       2563 GGGTATCACCTTCAAACTTGACTTC----------AGCACGTGTCTTGTAGTTCCCGTC 2515

Qy       1139 TGAAGCGGCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAA 1198
                |  |  | ||  || ||        |   |  |||| |  |  |  |||
Db       2514 ATCCTTGAAAAATATAGTTCTTTCCTGTACATAACCTTCGGGCATGGCACTCTTGAAAAA 2455

Qy       1199 ACGCACGTTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACT 1258
                |  |||     | || |  ||          |   | | |  || |
Db       2454 GTCATGCCGTTTCATATGATCTGGGTATCTTGAAAAGCATTGAACACCATAAGAGAAAGT 2395

Qy       1259 GTTCGTTCAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGAT 1318
                |  |   |  | || || |  | |  | | | | | | | | | | | | |
Db       2394 AGTGACAAGTGTTGGCCATGGAACAGGTAGTTTTCCAGTAGTGCAAATAAATTTAAGGGT 2335

Qy       1319 GGTCAGCCCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGC 1378
                |  || ||  | |  | |  | | | | | |  ||||
Db       2334 AAGTTTTC-----CGTATGTTGCATCACCTTCACCCTCTCCACTGACAGAAAATTTGTGC 2280

Qy       1379 CGGTGTGCAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTA 1438
                | |  ||  ||  |  |  ||||| |  |  |  | | |
Db       2279 CCATTAACATCACCATCTAATTCAACAAGAATTGGGACAACTCCAGTGAAAAGTTCTTCT 2220

Qy       1439 TCCTGGCTGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGC 1498
                | | ||  |  |  |  |  |  |  | | ||||  ||||||
Db       2219 CCTTTACTCATCGGTACCCAGCTTTTGTTCCCTTTAGTGAGGGTTA-----ATTGCGCGC 2165

Qy       1499 TTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCA 1558
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2164 TTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCA 2105

Qy       1559 CACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAA 1618
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2104 CACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAA 2045

Qy       1619 CTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAG 1678
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2044 CTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAG 1985

Qy       1679 CTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCC 1738
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1984 CTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCC 1925

Qy       1739 GCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCT 1798
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1924 GCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCT 1865

Qy       1799 CACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATG 1858
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1864 CACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATG 1805

Qy       1859 TGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTC 1918
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1804 TGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTC 1745

Qy       1919 CATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGA 1978
```

```
Db       1744 CATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGA 1685
Qy       1979 AACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCT 2038
Db       1684 AACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCT 1625
Qy       2039 CCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTG 2098
Db       1624 CCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTG 1565
Qy       2099 GCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAG 2158
Db       1564 GCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAG 1505
Qy       2159 CTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTAT 2218
Db       1504 CTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTAT 1445
Qy       2219 CGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAAC 2278
Db       1444 CGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAAC 1385
Qy       2279 AGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAAC 2338
Db       1384 AGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAAC 1325
Qy       2339 TACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTC 2398
Db       1324 TACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTC 1265
Qy       2399 GGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTT 2458
Db       1264 GGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTT 1205
Qy       2459 TTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATC 2518
Db       1204 TTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATC 1145
Qy       2519 TTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATG 2578
Db       1144 TTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATG 1085
Qy       2579 AGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCA 2638
Db       1084 AGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCA 1025
Qy       2639 ATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCA 2698
Db       1024 ATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCA 965
Qy       2699 CCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAG 2758
Db        964 CCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAG 905
Qy       2759 ATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGAC 2818
Db        904 ATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGAC 845
Qy       2819 CCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGC 2878
Db        844 CCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGC 785
Qy       2879 AGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCT 2938
Db        784 AGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCT 725
Qy       2939 AGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATC 2998
Db        724 AGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATC 665
Qy       2999 GTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGG 3058
```

```
Db        664 GTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGG 605
Qy       3059 CGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATC 3118
Db        604 CGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATC 545
Qy       3119 GTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAAT 3178
Db        544 GTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAAT 485
Qy       3179 TCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAG 3238
Db        484 TCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAG 425
Qy       3239 TCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGAT 3298
Db        424 TCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGAT 365
Qy       3299 AATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGG 3358
Db        364 AATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGG 305
Qy       3359 CGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCA 3418
Db        304 CGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCA 245
Qy       3419 CCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGA 3478
Db        244 CCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGA 185
Qy       3479 AGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTC 3538
Db        184 AGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTC 125
Qy       3539 TTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATA 3598
Db        124 TTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATA 65
Qy       3599 TTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTG 3658
Db         64 TTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTG 5
Qy       3659 CCAC 3662
Db          4 CCAC 1
```

RESULT 9
ADP26615/c
ID ADP26615 standard; DNA; 3637 BP.
XX
AC ADP26615;
XX
DT 26-AUG-2004 (first entry)
XX
DE Green fluorescent protein (GFP) plasmid DNA #1.
XX
KW Sequence variation; heteroduplex; transcription; DNA integration;
KW ribozyme expression; gene; ds; green fluorescent protein; GFP.
XX
OS Aequorea victoria.
OS Synthetic.
XX
PN US2004110130-A1.
XX
PD 10-JUN-2004.
XX
PF 25-OCT-2002; 2002US-00280913.
XX
PR 02-FEB-2001; 2001US-0266386P.
PR 14-FEB-2001; 2001US-0268785P.
PR 01-FEB-2002; 2002US-00066390.
PR 08-AUG-2002; 2002US-0402342P.

XX
PA (LARG-) LARGE SCALE BIOLOGY CORP.
XX
PI Padgett HS, Lindbo JA, Fitzmaurice WP;
XX
DR WPI; 2004-440326/41.
XX
PT Redistributing sequence variations between non-identical polynucleotide
PT sequences, useful for generating improved polynucleotide having a desired
PT characteristic, comprises making a heteroduplex and introducing a nick.
XX
PS Example 5; SEQ ID NO 31; 75pp; English.
XX
CC The invention relates to an in vitro method of redistributing sequence
CC variations between non-identical polynucleotide sequences, comprising
CC making a heteroduplex polynucleotide from two non-identical
CC polynucleotides, introducing a nick in the second strand at or near a
CC base pair mismatch site, removing the mismatched base(s) from the
CC mismatch site where the nick occurred and using the first strand as a
CC template to replace the removed base(s) with bases that complement the
CC base(s) in the first strand. The invention also relates to an in vitro
CC method of making a population of sequence variants from a heteroduplex
CC polynucleotide sequence, obtaining a polynucleotide sequence encoding a
CC desired functional property and identifying a reasserted DNA molecule
CC encoding a protein with a desired functional property. The method is
CC useful for generating an improved polynucleotide sequence or a population
CC of improved polynucleotide sequences possessing at least one desired
CC phenotypic characteristic (e.g., promotes transcription of linked
CC polynucleotides), where such polynucleotides are useful for expression
CC from a plant, animal, fungal, yeast, or bacterial expression vector, for
CC integration to form a transgenic plant, animal or microorganism, and for
CC expression of a ribozyme. This sequence represents DNA used in the scope
CC of the invention.
XX
SQ Sequence 3637 BP; 959 A; 879 C; 868 G; 931 T; 0 U; 0 Other;

Query Match 72.5%; Score 2656.4; DB 12; Length 3637;
Best Local Similarity 84.0%; Pred. No. 0;
Matches 3079; Conservative 0; Mismatches 556; Indels 29; Gaps 6;

```
Qy          1 CTAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3637 CTAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTC 3578

Qy         61 ATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3577 ATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGA 3518

Qy        121 GATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3517 GATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTC 3458

Qy        181 CAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3457 CAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACC 3398

Qy        241 CTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3397 CTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAG 3338

Qy        301 CCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3337 CCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAA 3278

Qy        361 AGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCAC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3277 AGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCAC 3218

Qy        421 CACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCCCATTCGCCATTCAGGCTGCG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3217 CACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCCCATTCGCCATTCAGGCTGCG 3158

Qy        481 CAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3157 CAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGG 3098

Qy        541 GGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3097 GGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTG 3038

Qy        601 TAAAACGACGGCCAGTGAGCGCGCCTCGTTCATTCACGTTTTTGAACCCGTGGAGGACGG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3037 TAAAACGACGGCCAGTGAGCGCGCGTAATACGACTCACTATAGGGCGAATTGGAGCTCCA 2978

Qy        661 GCAGACTCGCGGTGCAAATGTGTTTTACAGCGTGATGGAGCAGATGAAGATGCTCGACAC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2977 CCGCGGTGGCGGCCGCTCTAGAACTAGTGGATCCCCCGGGCTGCAGGAATTCTTATTTGT 2918

Qy        721 GCTGCAGAACACGCAGCTAGATTAACCCTAGAAAGATAATCATATTGTGACGTACGTTAA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2917 ATAGTTCATCCATGCCATGTGTAATCCCAGCAGCTGTTACAAACTCAAGAAGGACCATGT 2858

Qy        781 AGATAATC--ATGCGTAAAATTGACGCATGGGATCTGTAATACGACTCACTATAGGGCGA 838
              ||||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2857 GGTCTCTCTTTTCGTTGGGATCTTTCGAAAGGGCAGATTGTGTGGACAGGTAATGGTTGT 2798

Qy        839 ATTGGGTACCGGGCCCCCCCTCGAGGTCGACGGTATCGATAAGCTTGATATCGAATTCCT 898
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2797 CTGGTAAAAGGACAGGGCCATCGCCAATTGGAGTATTTTGTTGATAATGGTCTGCTAGTT 2738

Qy        899 GCAGCCCGGGGGATCCACTAGTTCTAGAGCGGCCGCCACCGCGGTGGAGCTCCAGCTTTT 958
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2737 GAACGCTTCCATCTTCAATGTTGTGTCTAATTTTGAAGTTAACTTTGATTCCATTCTTTT 2678

Qy        959 GTTCCCTTTAGTGAGGGTTAATTAGATCCCATGCGTCAATTTTACGCAGACTATCTTTCT 1018
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2677 GTTTGTCTGCCATGATGTATACATTGTGTGA-GTTATAGTTGTATTCCAATTTGTGTCCA 2619

Qy       1019 AGGGTTAATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCG 1078
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2618 AGAATGTTTCCATCTTCTTTAAAATCAATACCTTTTAACTCGATTC-----TATTAACAA 2564

Qy       1079 CCCAAGCTGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGT 1138
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2563 GGGTATCACCTTCAAACTTGACTTC----------AGCACGTGTCTTGTAGTTCCCGTC 2515

Qy       1139 TGAAGCGGCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAA 1198
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2514 ATCCTTGAAAAATATAGTTCTTTCCTGTACATAACCTTCGGGCATGGCACTCTTGAAAAA 2455

Qy       1199 ACGCACGTTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACT 1258
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2454 GTCATGCCGTTTCATATGATCTGGGTATCTTGAAAAGCATTGAACACCATAAGAGAAAGT 2395

Qy       1259 GTTCGTTCAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGAT 1318
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2394 AGTGACAAGTGTTGGCCATGGAACAGGTAGTTTTCCAGTAGTGCAAATAAATTTAAGGGT 2335

Qy       1319 GGTCAGCCCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGC 1378
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2334 AAGTTTTC-----CGTATGTTGCATCACCTTCACCCTCTCCACTGACAGAAAATTTGTGC 2280

Qy       1379 CGGTGTGCAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTA 1438
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2279 CCATTAACATCACCATCTAATTCAACAAGAATTGGGACAACTCCAGTGAAAAGTTCTTCT 2220

Qy       1439 TCCTGGCTGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGC 1498
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2219 CCTTTACTCATCGGTACCCAGCTTTTGTTCCCTTTAGTGAGGGTTA-----ATTGCGCGC 2165

Qy       1499 TTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCA 1558
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2164 TTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCA 2105

Qy       1559 CACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAA 1618
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       2104 CACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAA 2045
Qy       1619 CTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAG 1678
Db       2044 CTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAG 1985
Qy       1679 CTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCC 1738
Db       1984 CTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCC 1925
Qy       1739 GCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCT 1798
Db       1924 GCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCT 1865
Qy       1799 CACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATG 1858
Db       1864 CACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATG 1805
Qy       1859 TGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTC 1918
Db       1804 TGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTC 1745
Qy       1919 CATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGA 1978
Db       1744 CATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGA 1685
Qy       1979 AACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCT 2038
Db       1684 AACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCT 1625
Qy       2039 CCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTG 2098
Db       1624 CCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTG 1565
Qy       2099 GCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAG 2158
Db       1564 GCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAG 1505
Qy       2159 CTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTAT 2218
Db       1504 CTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTAT 1445
Qy       2219 CGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAAC 2278
Db       1444 CGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAAC 1385
Qy       2279 AGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAAC 2338
Db       1384 AGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAAC 1325
Qy       2339 TACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTC 2398
Db       1324 TACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTC 1265
Qy       2399 GGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTT 2458
Db       1264 GGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTT 1205
Qy       2459 TTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATC 2518
Db       1204 TTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATC 1145
Qy       2519 TTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATG 2578
Db       1144 TTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATG 1085
Qy       2579 AGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCA 2638
Db       1084 AGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCA 1025
Qy       2639 ATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCA 2698
Db       1024 ATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCA 965
```

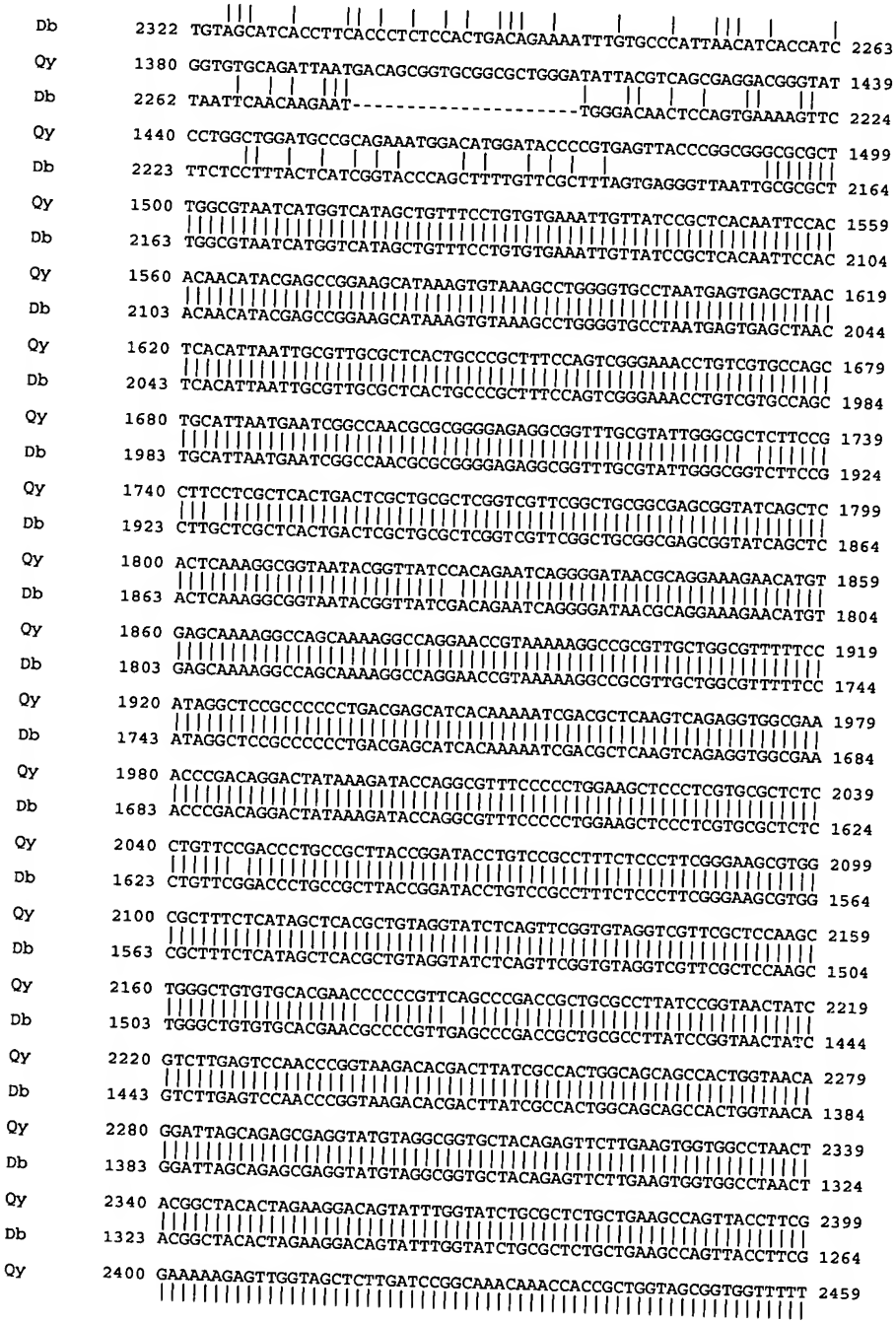
```
Qy       2699 CCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAG 2758
Db        964 CCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAG 905
Qy       2759 ATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGAC 2818
Db        904 ATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGAC 845
Qy       2819 CCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGC 2878
Db        844 CCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGC 785
Qy       2879 AGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCT 2938
Db        784 AGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCT 725
Qy       2939 AGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATC 2998
Db        724 AGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATC 665
Qy       2999 GTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGG 3058
Db        664 GTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGG 605
Qy       3059 CGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATC 3118
Db        604 CGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATC 545
Qy       3119 GTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAAT 3178
Db        544 GTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAAT 485
Qy       3179 TCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAG 3238
Db        484 TCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAG 425
Qy       3239 TCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGAT 3298
Db        424 TCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGAT 365
Qy       3299 AATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGG 3358
Db        364 AATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGG 305
Qy       3359 CGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCA 3418
Db        304 CGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCA 245
Qy       3419 CCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGA 3478
Db        244 CCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGA 185
Qy       3479 AGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTC 3538
Db        184 AGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTC 125
Qy       3539 TTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATA 3598
Db        124 TTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATA 65
Qy       3599 TTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTG 3658
Db          64 TTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTG 5
Qy       3659 CCAC 3662
Db           4 CCAC 1
```

RESULT 10
ADI26329/c
ID ADI26329 standard; DNA; 3637 BP.

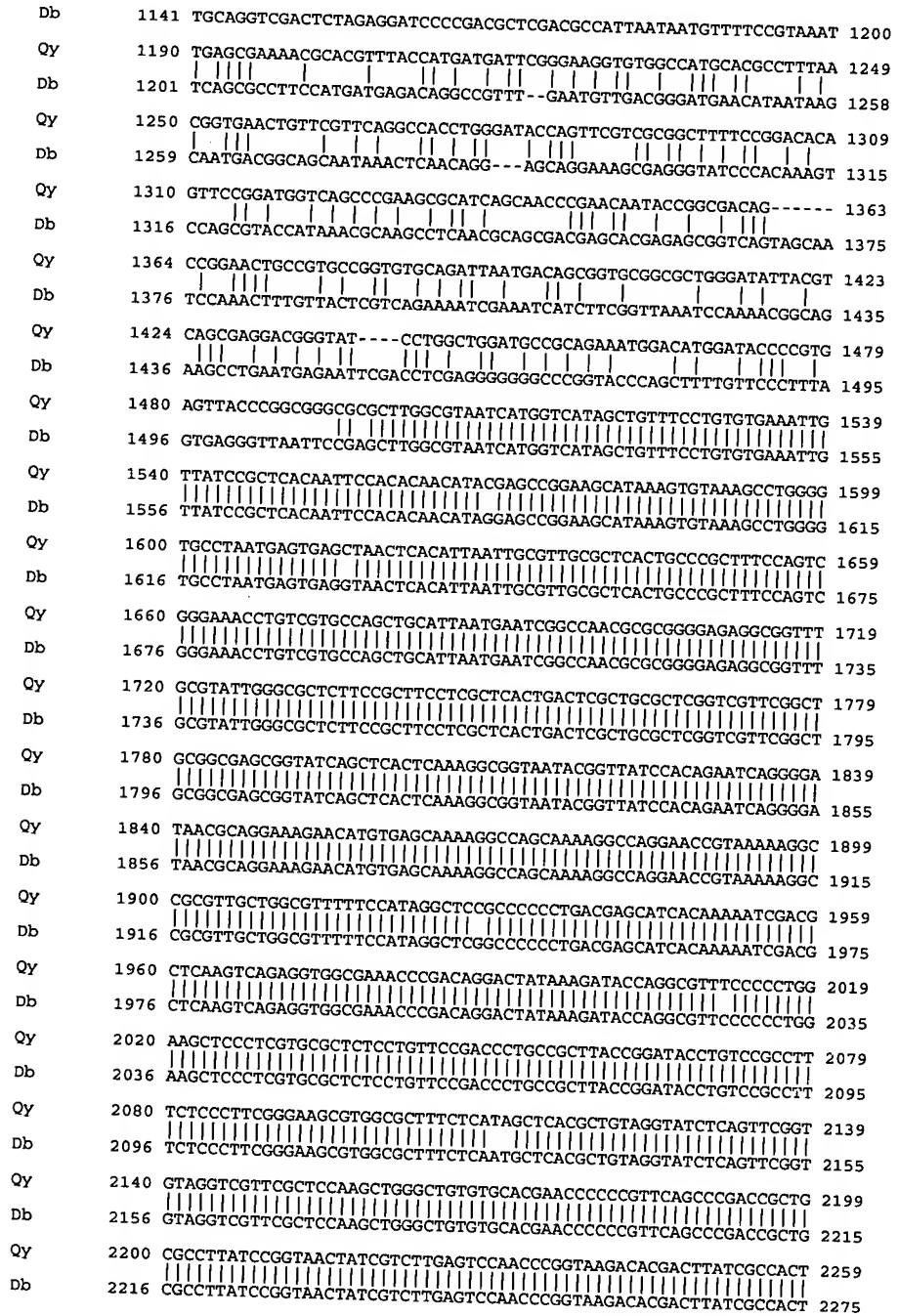
XX
AC ADI26329;
XX
DT 22-APR-2004 (first entry)
XX
DE Novel endonuclease Res I-related plasmid pBSC3BFP.
XX
KW endonuclease; molecular biology; plant propagation; phenotypic trait;
KW herbicide tolerance; heat tolerance; cold tolerance; drought; salinity;
KW osmotic stress; pest resistance; insect; nematode; arachnid; fungal;
KW bacterial; viral; enzyme production; secondary metabolite;
KW male sterility; female sterility; dwarfness; early maturity; Res I;
KW pBSC3BFP; ds.
XX
OS Unidentified.
XX
PN US2003148315-A1.
XX
PD 07-AUG-2003.
XX
PF 01-AUG-2002; 2002US-00211079.
XX
PR 01-FEB-2002; 2002US-0353722P.
PR 14-MAR-2002; 2002US-00098155.
XX
PA (PADG/) PADGETT H S.
PA (VAEW/) VAEWHONGS A A.
XX
PI Padgett HS, Vaewhongs AA;
XX
DR WPI; 2003-897548/82.
XX
PT New nucleic acid molecule encoding endonucleases, useful in molecular
PT biology, specifically to generating populations of related nucleic acid
PT molecules, and in plant propagation with useful phenotypic traits.
XX
PS Example 14; Fig 5; 46pp; English.
XX
CC This invention relates to a novel endonuclease (Res I) nucleic acid
CC molecule which comprises a fully defined sequence of 899 bp given in the
CC specification. The methods and compositions of the present invention are
CC useful in molecular biology, and more specifically to generating
CC populations of related nucleic acid molecules. They may also be used in
CC plant propagation with useful phenotypic traits, such as improved
CC tolerance to herbicides, improved tolerance to extremes of heat or cold,
CC drought, salinity or osmotic stress, improved resistance to pests
CC (insects, nematodes or arachnids) or diseases (fungal, bacterial or
CC viral), production of enzymes or secondary metabolites, male or female
CC sterility, dwarfness and early maturity. The present sequence is that of
CC a plasmid which is related to the invention.
XX
SQ Sequence 3637 BP; 942 A; 900 C; 860 G; 935 T; 0 U; 0 Other;

Query Match 72.3%; Score 2647.8; DB 10; Length 3637;
Best Local Similarity 83.4%; Pred. No. 0;
Matches 3054; Conservative 0; Mismatches 582; Indels 27; Gaps 3;

```
Qy          1 CTAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTC  60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3637 CTAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTC 3578

Qy         61 ATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGA 120
              ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db       3577 ATTTTTTAACCAATAAGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGA 3518

Qy        121 GATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTC 180
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db       3517 GATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCAGTATTAAAGAACGTGGACTC 3458

Qy        181 CAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3457 CAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACC 3398

Qy        241 CTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3397 CTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAG 3338

Qy        301 CCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3337 CCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAA 3278

Qy        361 AGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCAC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3277 AGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCAC 3218

Qy        421 CACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCCCATTCGCCATTCAGGCTGCG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3217 CACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCCCATTCGCCATTCAGGCTGCG 3158

Qy        481 CAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3157 CAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGG 3098

Qy        541 GGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3097 GGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTG 3038

Qy        601 TAAAACGACGGCCAGTGAGCGCGCCTCGTTCATTCACGTTTTTGAACCCGTGGAGGACGG 660
              ||||||||||||||||||||| | | |  | | |  |||| |
Db       3037 TAAAACGACGGCCAGTGAGCGCGCGTAATACGACTCACTATAGGGCGAATTGGAGCTCCA 2978

Qy        661 GCAGACTCGCGGTGCAAATGTGTTTTACAGCGTGATGGAGCAGATGAAGATGCTCGACAC 720
              | | ||||  |  |  |  | | | | | | | |
Db       2977 CCGCGGTGGCGGCCGCTCTAGAACTAGTGGATCCCCCGGGCTGCAGGAATTCTTATTTGT 2918

Qy        721 GCTGCAGAACACGCAGCTAGATTAACCCTAGAAAGATAATCATATTGTGACGTACGTTAA 780
              | | | |  | | |||  | | | | |  || | ||  |
Db       2917 ATAGTTCATCCATGCCATGTGTAATCCCAGCAGCAGTTACAAACTCAAGAAGGACCATGT 2858

Qy        781 AGATAATC-ATGCGTAAAATTGACGCATGGGATCTGTAATACGACTCACTATAGGGCGAA 839
              | | | | ||| |  |  | | | |  ||  | |
Db       2857 GGTCACGCTTTTCGTTGGGATCTTTCGAAAGGGCAGATTGTGTCGACAGGTAATGGTTGT 2798

Qy        840 TTGGGTACCGGGCCCCCCCTCGAGGTCGACGGTATCGATAAGCTTGATATCGAATTCCTG 899
              ||| | || |  ||  |  |  | || | |  |
Db       2797 CTGGTAAAAGGACAGGGCCATCGCCAATTGGAGTATTTTGTTGATAATGGTCTGCTAGTT 2738

Qy        900 CAGCCCGGGGGATCCACTAGTTCTAGAGCGGCCGCCACCGCGGTGGAGCTCCAGCTTTTG 959
              | |  |||  | | |  |  | |  |||| |||
Db       2737 GAACGGATCCATCTTCAATGTTGTGGCGAATTTTGAAGTTAGCTTTGATTCCATTCTTTT 2678

Qy        960 TTCCCTTTAGTGAGGGTTAATTAGATCCCATGCGTCAATTTTACGCAGACTATCTTTCTA 1019
              | | | | |  ||  |  |  | || ||| |  |  | | | |
Db       2677 GTTTGTCTGCCGTGATGTATACATTGTGTGAGTTAAAGTTGTACTCGAGTTTGTGTCCGA 2618

Qy       1020 GGGTTAATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGC 1079
              | |  || | | | | | | ||||  | |  ||  | |  | | |
Db       2617 GAATGTTTCCATCTTCTTTAAAATCAATACCTTTTAACTCGATAC-----GATTAACAAG 2563

Qy       1080 CCAAGCTGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTT 1139
              | |  | | || | | ||  |  |  | ||| |  |  | ||
Db       2562 GGTATCACCTTCAAACTTGACTTCAGCACGCGTCTTGTAGTTCCCGTCATCTTTGAAAGA 2503

Qy       1140 GAAGCGGCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAA 1199
              |  || |  |  | || |  |  || |  | | |  || |
Db       2502 TATAGTGCGTTCCTGTACATAACCTTCGGGCATGGCACTCTTGAAAAAGTCATGCCGTTT 2443

Qy       1200 CGCACGTTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTG 1259
              | | | |  || | ||  | |  ||||  |  |  | |  | |
Db       2442 CATATGATCCGGATAACGAGAAAAGCATTGAACACCATGAGAGAAAGTAGTGACAAGTGT 2383

Qy       1260 TTCGTTCAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATG 1319
              |  | || | | | |||||  |  || |  || | |
Db       2382 TGGCCATGGAACAGGTAGTTTTCCAGTAGTGCAAATAAATTTAAGTGTAAGCTTTCCGTA 2323

Qy       1320 GTCAGCCCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCC 1379
```


Db 2322 TGTAGCATCACCTTCACCCTCTCCACTGACAGAAAATTTGTGCCCATTAACATCACCATC 2263
Qy 1380 GGTGTGCAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTAT 1439
Db 2262 TAATTCAACAAGAAT----------TGGGACAACTCCAGTGAAAAGTTC 2224
Qy 1440 CCTGGCTGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCT 1499
Db 2223 TTCTCCTTTACTCATCGGTACCCAGCTTTTGTTCGCTTTAGTGAGGGTTAATTGCGCGCT 2164
Qy 1500 TGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCAC 1559
Db 2163 TGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCAC 2104
Qy 1560 ACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAAC 1619
Db 2103 ACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAAC 2044
Qy 1620 TCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGC 1679
Db 2043 TCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGC 1984
Qy 1680 TGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCG 1739
Db 1983 TGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGGTCTTCCG 1924
Qy 1740 CTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTC 1799
Db 1923 CTTGCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTC 1864
Qy 1800 ACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGT 1859
Db 1863 ACTCAAAGGCGGTAATACGGTTATCGACAGAATCAGGGGATAACGCAGGAAAGAACATGT 1804
Qy 1860 GAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCC 1919
Db 1803 GAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCC 1744
Qy 1920 ATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAA 1979
Db 1743 ATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAA 1684
Qy 1980 ACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTC 2039
Db 1683 ACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTC 1624
Qy 2040 CTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGG 2099
Db 1623 CTGTTCGGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGG 1564
Qy 2100 CGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGC 2159
Db 1563 CGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGC 1504
Qy 2160 TGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATC 2219
Db 1503 TGGGCTGTGTGCACGAACGCCCCGTTGAGCCCGACCGCTGCGCCTTATCCGGTAACTATC 1444
Qy 2220 GTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACA 2279
Db 1443 GTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACA 1384
Qy 2280 GGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACT 2339
Db 1383 GGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACT 1324
Qy 2340 ACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCG 2399
Db 1323 ACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCG 1264
Qy 2400 GAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTT 2459
Db 1263 GAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTT 1204
Qy 2460 TTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCT 2519
Db 1203 TTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCT 1144
Qy 2520 TTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGA 2579
Db 1143 TTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGA 1084
Qy 2580 GATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAA 2639
Db 1083 GATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAA 1024
Qy 2640 TCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCAC 2699
Db 1023 TCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCAC 964
Qy 2700 CTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGA 2759
Db 963 CTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGA 904
Qy 2760 TAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACC 2819
Db 903 TAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACC 844
Qy 2820 CACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCA 2879
Db 843 CACGCTGACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCA 784
Qy 2880 GAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTA 2939
Db 783 GAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTA 724
Qy 2940 GAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCG 2999
Db 723 GAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCG 664
Qy 3000 TGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGC 3059
Db 663 TGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGC 604
Qy 3060 GAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCG 3119
Db 603 GAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCG 544
Qy 3120 TTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATT 3179
Db 543 TTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATT 484
Qy 3180 CTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGT 3239
Db 483 CTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGT 424
Qy 3240 CATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATA 3299
Db 423 CATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATA 364
Qy 3300 ATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGC 3359
Db 363 ATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGG 304
Qy 3360 GAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCAC 3419
Db 303 GAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCAC 244
Qy 3420 CCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAA 3479
Db 243 CCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAA 184
Qy 3480 GGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCT 3539
Db 183 GGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCT 124

```
Qy       3540 TCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATAT 3599
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 TCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATAT 64

Qy       3600 TTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGC 3659
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 TTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGC 4

Qy       3660 CAC 3662
             |||
Db          3 CAC 1


RESULT 11
AAQ13578
ID   AAQ13578 standard; DNA; 3681 BP.
XX
AC   AAQ13578;
XX
DT   02-DEC-1991  (first entry)
XX
DE   Plasmid pKSEL5.
XX
KW   recombinant expression plasmid; phage phi X174; protein E;
KW   bacterial ghost; ss.
XX
OS   Synthetic.
XX
PN   WO9113155-A.
XX
PD   05-SEP-1991.
XX
PF   24-FEB-1990;   90DE-04005874.
XX
PR   24-FEB-1990;   90DE-04005874.
XX
PA   (BOEF ) BOEHRINGER MANNHEIM GMBH.
XX
PI   Lubitz W,  Szostak MP;
XX
DR   WPI; 1991-281471/38.
XX
PT   Carrier recombinant protein for vaccines against HIV, etc. - obtd. by
PT   expressing fusion protein gene in gram negative bacteria and gene that
PT   encodes for lytic membrane protein.
XX
PS   Example 1; Page 33-34; 45pp; German.
XX
CC   Plasmid pKSEL5 comprises three multiple cloning sites (mcs1-3), the
CC   ampicilline resistance gene, regions of the lac operon and a partial
CC   phage phi X174 E' target sequence. This plasmid can be used as a carrier
CC   vector according to the invention for the insertion of sequences encoding
CC   viral antigenic peptides, e.g. the HIV gp41 sequence was isolated as a
CC   HincII/PvuII fragment from plasmid pHF14. The fragment also included a
CC   linker sequence and the last 45 codons of gp120. The fragment was ligated
CC   to pKSEL5 which had been cleaved with AccI and end-filled. Alternatively,
CC   a new carrier vector, denoted pMTV1 (see AAP13576), can be constructed
CC   from pKSEL5 which has been partially digested with SspI. It is ligated to
CC   a SspI-DraII fragment from pML1 (see AAQ13577). Foreign sequences are
CC   cloned into mcs2
XX
SQ   Sequence 3681 BP; 949 A; 917 C; 932 G; 883 T; 0 U; 0 Other;

  Query Match             72.0%;  Score 2635.8;  DB 2;  Length 3681;
  Best Local Similarity   83.4%;  Pred. No. 0;
  Matches 3073;  Conservative    0;  Mismatches  582;  Indels   28;  Gaps    6;

Qy          3 AAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTCAT 62
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AAATTGTAAACGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTCAT 60

Qy         63 TTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGAGA 122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGAGA 120

Qy        123 TAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTCCA 182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTCCA 180

Qy        183 ACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCCT 242
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCCT 240

Qy        243 AATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAGCC 302
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAGCC 300

Qy        303 CCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAAAG 362
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAAAG 360

Qy        363 CGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCACCA 422
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCACCA 420

Qy        423 CACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCCCATTCGCCATTCAGGCTGCGCA 482
              |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db        421 CACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCCCATTCGCCATTCAGGCTACGCA 480

Qy        483 ACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGGG 542
              |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db        481 ACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAGGGGG 540

Qy        543 GATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTGTA 602
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTGTA 600

Qy        603 AAACGACGGCCAGTGAGCGCGCCTCGTTCATTCACGTTTTTGAACCCGTGGAGGACGGGC 662
              |||||||||||||||     |   |   |     |||   | |    || |
Db        601 AAACGACGGCCAGTGAATTGTAATACGACTCACTATAGGGCGAATTGGAGCTCCACCGCG 660

Qy        663 AGACTCGCGGTGCAAATGTGTTTTACAGCGTGATGGAGCAGATGAAGATGCTCGACACGC 722
                 || |  | | | || |  ||     |  | | | |        | |
Db        661 GTGGCGGCCGCTCTAGTATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTA 720

Qy        723 TGCAGAACACGCAGCTAGATTAACCCTAGAAAGATAATCATATTGTGACGTACGTTAAAG 782
              ||  |||          ||    || | |||       |          ||   || |
Db        721 AGCCAGTATATACACTCCGCTATCGCTACGTGACTGGGTCATGGCTGCGCCCCGACACCC 780

Qy        783 ATAATCATGCGTAAAATTGACGCATGGGATCTGTAATACGACTCACTATAGGGCGAATTG 842
                | ||  ||   |  | |      ||  |||       |   |       |
Db        781 GCCAACACCCGCTGACGCGCCCTGACGGGCTTGTCTGCTCCCGGCATCCGCTTACAGACA 840

Qy        843 GGTACCGGGCCCCCCCCTCGAGGTCGACGGTATCGA------TAAGCTTGATATCGAATTC 896
              |   |  | | || ||| | | |    ||      |   |   | ||||
Db        841 AGCTGTGACCGTCTCCGGGAGCTGCATGTGTCAGAGGTTTTCACCGTCATCACCGAAACG 900

Qy        897 CTGCAGCCCGGGGGATCCACTAGTTCTAGAGCGGCCGCCACCGCGGTGGAGCTCCAGCTT 956
              |   || | |  | || |  || |||| ||      |||   |||   |||
Db        901 CGCGAGGCAGTAAGGTCGGATGCTTTGTGAGCAATTCGTCCCTTAAGTAAGCAATTGCTG 960

Qy        957 TTGTTCCCTTTAGTGAGGGTTAATTAGATCCCAT-------GCGTCAATTTTACGCAGAC 1009
              |   | |      || | |    |||  |     |  |||||   |  |
Db        961 TAAAGTCGTCACTGTGCGGATCACCGCTTCCAGTAGCGACAGAAGCAATTGATTGGTAAA 1020

Qy       1010 TATCTTTCTAGGGTTAATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCT 1069
              | ||   | | |   | | |  | |  | | || | |  | |||  | ||
Db       1021 TTTCGAGAGAAAGATCGCGAGGAAGATCAATACATAAAGAGTTGAACTTCTTTGTTGTCT 1080

Qy       1070 CAGTCATCGCCCAAGCTGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTC 1129
              | ||| |  | || | |  | |||  | |  |  |  |  | | | |
Db       1081 TCGACATGGGTAATCCTCATGTTTGAATGGCCCTAGAGGATCCGGCCAAGCTTGCATGCC 1140

Qy       1130 CCGCGAGGTTGAAGCGGCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGT 1189
                 |       |    ||   |  || | |  || |    |  |
```

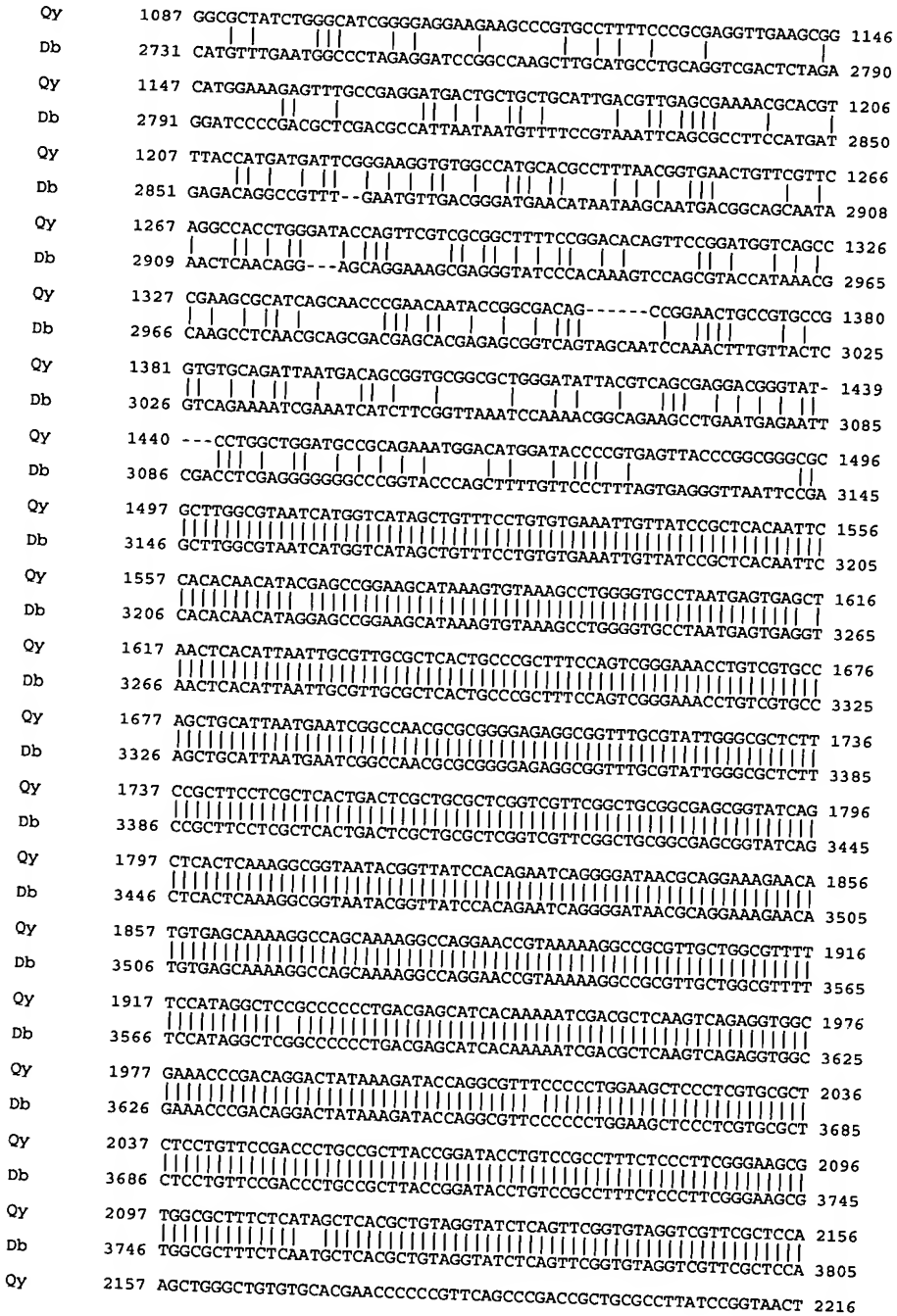

Db 1141 TGCAGGTCGACTCTAGAGGATCCCCGACGCTCGACGCCATTAATAATGTTTTCCGTAAAT 1200
Qy 1190 TGAGCGAAAACGCACGTTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAA 1249
Db 1201 TCAGCGCCTTCCATGATGAGACAGGCCGTTT--GAATGTTGACGGGATGAACATAATAAG 1258
Qy 1250 CGGTGAACTGTTCGTTCAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACA 1309
Db 1259 CAATGACGGCAGCAATAAACTCAACAGG---AGCAGGAAAGCGAGGGTATCCCACAAAGT 1315
Qy 1310 GTTCCGGATGGTCAGCCCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAG------ 1363
Db 1316 CCAGCGTACCATAAACGCAAGCCTCAACGCAGCGACGAGCACGAGAGCGGTCAGTAGCAA 1375
Qy 1364 CCGGAACTGCCGTGCCGGTGTGCAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGT 1423
Db 1376 TCCAAACTTTGTTACTCGTCAGAAAATCGAAATCATCTTCGGTTAAATCCAAAACGGCAG 1435
Qy 1424 CAGCGAGGACGGGTAT----CCTGGCTGGATGCCGCAGAAATGGACATGGATACCCCGTG 1479
Db 1436 AAGCCTGAATGAGAATTCGACCTCGAGGGGGGGCCCGGTACCCAGCTTTTGTTCCCTTTA 1495
Qy 1480 AGTTACCCGGCGGGCGCGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTG 1539
Db 1496 GTGAGGGTTAATTCCGAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTG 1555
Qy 1540 TTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGG 1599
Db 1556 TTATCCGCTCACAATTCCACACAACATAGGAGCCGGAAGCATAAAGTGTAAAGCCTGGGG 1615
Qy 1600 TGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTC 1659
Db 1616 TGCCTAATGAGTGAGGTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTC 1675
Qy 1660 GGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTT 1719
Db 1676 GGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTT 1735
Qy 1720 GCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCT 1779
Db 1736 GCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCT 1795
Qy 1780 GCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGA 1839
Db 1796 GCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGA 1855
Qy 1840 TAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGC 1899
Db 1856 TAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGC 1915
Qy 1900 CGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACG 1959
Db 1916 CGCGTTGCTGGCGTTTTTCCATAGGCTCGGCCCCCCTGACGAGCATCACAAAAATCGACG 1975
Qy 1960 CTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGG 2019
Db 1976 CTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTCCCCCCTGG 2035
Qy 2020 AAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTT 2079
Db 2036 AAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTT 2095
Qy 2080 TCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGT 2139
Db 2096 TCTCCCTTCGGGAAGCGTGGCGCTTTCTCAATGCTCACGCTGTAGGTATCTCAGTTCGGT 2155
Qy 2140 GTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTG 2199
Db 2156 GTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTG 2215
Qy 2200 CGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACT 2259
Db 2216 CGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACT 2275
Qy 2260 GGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTT 2319
Db 2276 GGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTT 2335
Qy 2320 CTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCT 2379
Db 2336 CTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCT 2395
Qy 2380 GCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCAC 2439
Db 2396 GCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCAC 2455
Qy 2440 CGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATC 2499
Db 2456 CGTCGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATC 2515
Qy 2500 TCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACG 2559
Db 2516 TCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACG 2575
Qy 2560 TTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTA 2619
Db 2576 TTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTA 2635
Qy 2620 AAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCA 2679
Db 2636 AAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCA 2695
Qy 2680 ATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGC 2739
Db 2696 ATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGC 2755
Qy 2740 CTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGC 2799
Db 2756 CTGACTGCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGC 2815
Qy 2800 TGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCC 2859
Db 2816 TGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCC 2875
Qy 2860 AGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTAT 2919
Db 2876 AGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTAT 2935
Qy 2920 TAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGT 2979
Db 2936 TAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGT 2995
Qy 2980 TGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTC 3039
Db 2996 TGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTC 3055
Qy 3040 CGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAG 3099
Db 3056 CGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGAAAAAAAGCGGTTAG 3115
Qy 3100 CTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGT 3159
Db 3116 CTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGCT 3175
Qy 3160 TATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGAC 3219
Db 3176 TATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGAC 3235
Qy 3220 TGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTG 3279
Db 3236 TGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTG 3295
Qy 3280 CCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCAT 3339
Db 3296 CCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCAT 3355

```
Qy       3340 TGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTC 3399
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3356 TGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTC 3415

Qy       3400 GATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTC 3459
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3416 GATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTC 3475

Qy       3460 TGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAA 3519
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3476 TGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAA 3535

Qy       3520 ATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTG 3579
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3536 ATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTG 3595

Qy       3580 TCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCG 3639
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3596 TCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCG 3655

Qy       3640 CACATTTCCCCGAAAAGTGCCAC 3662
             |||||||||||||||||||||||
Db       3656 CACATTTCCCCGAAAAGTGCCAC 3678
```

RESULT 12
AAQ13576
ID AAQ13576 standard; DNA; 5314 BP.
XX
AC AAQ13576;
XX
DT 02-DEC-1991 (first entry)
XX
DE Plasmid pMTV1 containing 3 multiple cloning sites.
XX
KW recombinant expression plasmid; phage phi X174; protein E;
KW bacterial ghost; ss.
XX
OS Synthetic.
XX
PN WO9113155-A.
XX
PD 05-SEP-1991.
XX
PF 24-FEB-1990; 90DE-04005874.
XX
PR 24-FEB-1990; 90DE-04005874.
XX
PA (BOEF) BOEHRINGER MANNHEIM GMBH.
XX
PI Lubitz W, Szostak MP;
XX
DR WPI; 1991-281471/38.
XX
PT Carrier recombinant protein for vaccines against HIV, etc. - obtd. by
PT expressing fusion protein gene in gram negative bacteria and gene that
PT encodes for lytic membrane protein.
XX
PS Example 10; Page 28-29; 45pp; German.
XX
CC The carrier plasmid pMTV1 comprises the phage phi X174 E protein gene,
CC the ampicilline resistance gene, the lambda cI857 repressor gene and P(R)
CC from the lambda promoter/operator system. There are also lac operon
CC sequences and three multiple cloning sites. A foreign gene can be
CC inserted into the second multiple cloning site (mcs2). The protein
CC encoded by the foreign gene is especially a viral antigenic peptide, e.g.
CC of HIV, HBV or EBV. See also AAQ13577 and AAQ13578
XX
SQ Sequence 5314 BP; 1438 A; 1276 C; 1310 G; 1290 T; 0 U; 0 Other;

Query Match 71.6%; Score 2622; DB 2; Length 5314;
Best Local Similarity 83.4%; Pred. No. 0;
Matches 3058; Conservative 0; Mismatches 580; Indels 28; Gaps 6;

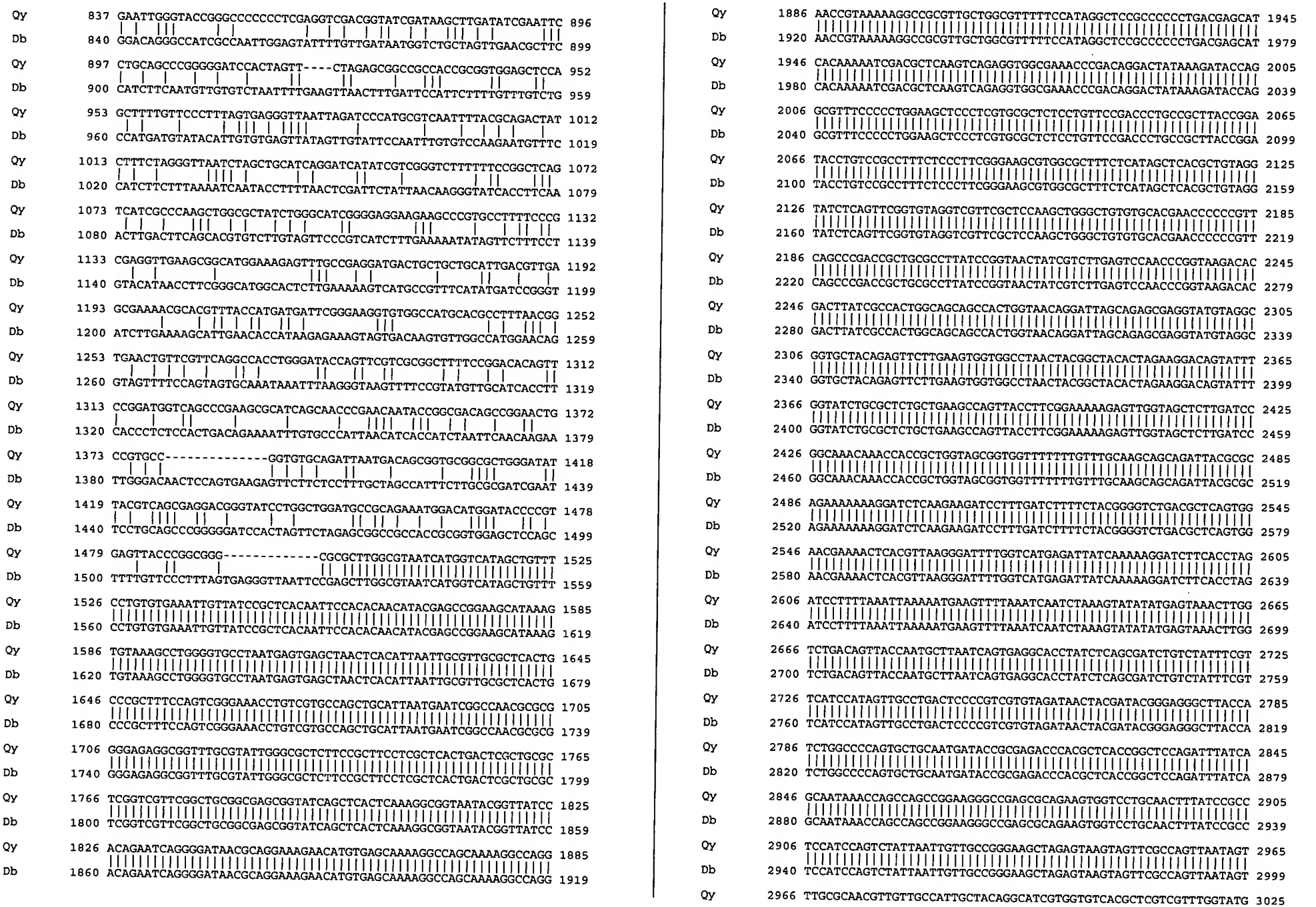
```
Qy         20 ATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTCATTTTTTAACCAATAGGCC 79
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1651 ATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTCATTTTTTAACCAATAGGCC 1710

Qy         80 GAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGAGATAGGGTTGAGTGTTGTT 139
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1711 GAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGAGATAGGGTTGAGTGTTGTT 1770

Qy        140 CCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTCCAACGTCAAAGGGCGAAAA 199
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1771 CCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTCCAACGTCAAAGGGCGAAAA 1830

Qy        200 ACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCCTAATCAAGTTTTTTGGGG 259
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1831 ACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCCTAATCAAGTTTTTTGGGG 1890

Qy        260 TCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAGCCCCCGATTTAGAGCTTGA 319
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1891 TCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAGCCCCCGATTTAGAGCTTGA 1950

Qy        320 CGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAAAGCGAAAGGAGCGGGCGCT 379
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1951 CGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAAAGCGAAAGGAGCGGGCGCT 2010

Qy        380 AGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCACCACACCCGCCGCGCTTAAT 439
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2011 AGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCACCACACCCGCCGCGCTTAAT 2070

Qy        440 GCGCCGCTACAGGGCGCGTCCCATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGGGCGA 499
              ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db       2071 GCGCCGCTACAGGGCGCGTCCCATTCGCCATTCAGGCTACGCAACTGTTGGGAAGGGCGA 2130

Qy        500 TCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGGGGATGTGCTGCAAGGCGA 559
              ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db       2131 TCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAGGGGGGATGTGCTGCAAGGCGA 2190

Qy        560 TTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTGTAAAACGACGGCCAGTGAG 619
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2191 TTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTGTAAAACGACGGCCAGTGAA 2250

Qy        620 CGCGCCTCGTTCATTCACGTTTTTGAACCCGTGGAGGACGGGCAGACTCGCGGTGCAAAT 679
                  |   |  |      |||  | |   || |      | || || | | |
Db       2251 TTGTAATACGACTCACTATAGGGCGAATTGGAGCTCCACCGCGGTGGCGGCCGCTCTAGT 2310

Qy        680 GTGTTTTACAGCGTGATGGAGCAGATGAAGATGCTCGACACGCTGCAGAACACGCAGCTA 739
               || |  ||   |  | | | |  | |    ||      ||
Db       2311 ATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGTATATACACTC 2370

Qy        740 GATTAACCCTAGAAAGATAATCATATTGTGACGTACGTTAAAGATAATCATGCGTAAAAT 799
               || | |||      |          ||    ||  |      | ||  ||   |
Db       2371 CGCTATCGCTACGTGACTGGGTCATGGCTGCGCCCCGACACCCGCCAACACCCGCTGACG 2430

Qy        800 TGACGCATGGGATCTGTAATACGACTCACTATAGGGCGAATTGGGTACCGGGCCCCCCCT 859
               | |    ||  |||      |  |      |    |   | | | ||
Db       2431 CGCCCTGACGGGCTTGTCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCG 2490

Qy        860 CGAGGTCGACGGTATCGA------TAAGCTTGATATCGAATTCCTGCAGCCCGGGGGATC 913
               ||| |  | |    ||      |   |   | ||||   |  || | |  | ||
Db       2491 GGAGCTGCATGTGTCAGAGGTTTTCACCGTCATCACCGAAACGCGCGAGGCAGTAAGGTC 2550

Qy        914 CACTAGTTCTAGAGCGGCCGCCACCGCGGTGGAGCTCCAGCTTTTGTTCCCTTTAGTGAG 973
               |  ||  ||||      ||      |||   ||| | |     | |
Db       2551 GGATGCTTTGTGAGCAATTCGTCCCTTAAGTAAGCAATTGCTGTAAAGTCGTCACTGTGC 2610

Qy        974 GGTTAATTAGATCCCAT-------GCGTCAATTTTACGCAGACTATCTTTCTAGGGTTAA 1026
               || | |    ||| |       |  |||||   |  | | ||     | | |
Db       2611 GGATCACCGCTTCCAGTAGCGACAGAAGCAATTGATTGGTAAATTTCGAGAGAAAGATCG 2670

Qy       1027 TCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGCT 1086
               | |  |  | | || | |  | |||  | ||  | ||| |  |  ||
Db       2671 CGAGGAAGATCAATACATAAAGAGTTGAACTTCTTTGTTGTCTTCGACATGGGTAATCCT 2730
```


Qy 1087 GGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCGG 1146
Db 2731 CATGTTTGAATGGCCCTAGAGGATCCGGCCAAGCTTGCATGCCTGCAGGTCGACTCTAGA 2790
Qy 1147 CATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACGT 1206
Db 2791 GGATCCCCGACGCTCGACGCCATTAATAATGTTTTCCGTAAATTCAGCGCCTTCCATGAT 2850
Qy 1207 TTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTTC 1266
Db 2851 GAGACAGGCCGTTT--GAATGTTGACGGGATGAACATAATAAGCAATGACGGCAGCAATA 2908
Qy 1267 AGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGCC 1326
Db 2909 AACTCAACAGG---AGCAGGAAAGCGAGGGTATCCCACAAAGTCCAGCGTACCATAAACG 2965
Qy 1327 CGAAGCGCATCAGCAACCCGAACAATACCGGCGACAG------CCGGAACTGCCGTGCCG 1380
Db 2966 CAAGCCTCAACGCAGCGACGAGCACGAGAGCGGTCAGTAGCAATCCAAACTTTGTTACTC 3025
Qy 1381 GTGTGCAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTAT- 1439
Db 3026 GTCAGAAAATCGAAATCATCTTCGGTTAAATCCAAAACGGCAGAAGCCTGAATGAGAATT 3085
Qy 1440 ---CCTGGCTGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGC 1496
Db 3086 CGACCTCGAGGGGGGGCCCGGTACCCAGCTTTTGTTCCCTTTAGTGAGGGTTAATTCCGA 3145
Qy 1497 GCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTC 1556
Db 3146 GCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTC 3205
Qy 1557 CACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCT 1616
Db 3206 CACACAACATAGGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGGT 3265
Qy 1617 AACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCC 1676
Db 3266 AACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCC 3325
Qy 1677 AGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTT 1736
Db 3326 AGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTT 3385
Qy 1737 CCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAG 1796
Db 3386 CCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAG 3445
Qy 1797 CTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACA 1856
Db 3446 CTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACA 3505
Qy 1857 TGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTT 1916
Db 3506 TGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTT 3565
Qy 1917 TCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGC 1976
Db 3566 TCCATAGGCTCGGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGC 3625
Qy 1977 GAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCT 2036
Db 3626 GAAACCCGACAGGACTATAAAGATACCAGGCGTTCCCCCCTGGAAGCTCCCTCGTGCGCT 3685
Qy 2037 CTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCG 2096
Db 3686 CTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCG 3745
Qy 2097 TGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCA 2156
Db 3746 TGGCGCTTTCTCAATGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCA 3805
Qy 2157 AGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACT 2216
Db 3806 AGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACT 3865
Qy 2217 ATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTA 2276
Db 3866 ATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTA 3925
Qy 2277 ACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTA 2336
Db 3926 ACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTA 3985
Qy 2337 ACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCT 2396
Db 3986 ACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCT 4045
Qy 2397 TCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTT 2456
Db 4046 TCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTT 4105
Qy 2457 TTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGA 2516
Db 4106 TTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGA 4165
Qy 2517 TCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCA 2576
Db 4166 TCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCA 4225
Qy 2577 TGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAAT 2636
Db 4226 TGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAAT 4285
Qy 2637 CAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGG 2696
Db 4286 CAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGG 4345
Qy 2697 CACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGT 2756
Db 4346 CACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTGCCCGTCGTGT 4405
Qy 2757 AGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAG 2816
Db 4406 AGATAACTACGATACGGGAGGGCTTACGATCTGGCCCCAGTGCTGCAATGATACCGCGAG 4465
Qy 2817 ACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGC 2876
Db 4466 ACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGC 4525
Qy 2877 GCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAG 2936
Db 4526 GCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAG 4585
Qy 2937 CTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCA 2996
Db 4586 CTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCA 4645
Qy 2997 TCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAA 3056
Db 4646 TCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAA 4705
Qy 3057 GGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGA 3116
Db 4706 GGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGA 4765
Qy 3117 TCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATA 3176
Db 4766 TCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGCTTATGGCAGCACTGCATA 4825
Qy 3177 ATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCA 3236
Db 4826 ATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCA 4885
Qy 3237 AGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGG 3296

```
Db       4886 AGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGG 4945

Qy       3297 ATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGG 3356
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4946 ATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGG 5005

Qy       3357 GGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTG 3416
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5006 GGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTG 5065

Qy       3417 CACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAG 3476
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5066 CACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAG 5125

Qy       3477 GAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATAC 3536
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5126 GAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATAC 5185

Qy       3537 TCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACA 3596
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5186 TCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACA 5245

Qy       3597 TATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAG 3656
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5246 TATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAG 5305

Qy       3657 TGCCAC 3662
             ||||||
Db       5306 TGCCAC 5311
```

RESULT 13
AAV14340
ID AAV14340 standard; DNA; 3699 BP.
XX
AC AAV14340;
XX
DT 19-MAY-1998 (first entry)
XX
DE Plasmid pBSGFP expressing green fluorescent protein.
XX
KW Green fluorescent protein; GFP; blue fluorescent protein; BFP; jellyfish;
KW genetic engineering marker; gene therapy; plasmid; ss.
XX
OS Synthetic.
XX
PN WO9742320-A1.
XX
PD 13-NOV-1997.
XX
PF 07-MAY-1997; 97WO-US007625.
XX
PR 08-MAY-1996; 96US-00646538.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Pavlakis GN, Gaitanaris GA, Stauber RH, Vournakis JN;
XX
DR WPI; 1997-558982/51.
XX
PT New nucleic acid encoding proteins of Aequorea victoria with increased
PT fluorescence - useful as markers for detecting cellular transformation,
PT subcellular localisation of proteins, for assessing gene therapy,
PT mutagenicity etc.
XX
PS Example; Page 67-68; 105pp; English.
XX
CC This sequence is a plasmid used to express the mutated green fluorescent
CC protein (GFP) of Aequorea victoria coding sequence of the invention. The
CC DNA of the invention that encodes a protein that is (a) a protein that
CC has Leu at position 65 (and optionally Thr at 168 and optionally further
CC Cys at 66) and has cellular fluorescence at least 5 times that of wild-
CC type GFP or (b) is a blue fluorescent protein (BFP) with His at position
CC 67 and also at least one of Leu at 65 and Ala at 164, and has cellular
CC fluorescence at least 5 times that of BFP (Tyr67 to His). The nucleic
CC acids can be used as markers in genetic engineering and gene therapy.
CC They may also be used to detect and characterise regulatory and coding
CC sequence elements that control subcellular expression and targeting of
CC proteins. Typical applications are monitoring targeting and transport of
CC proteins in cells; assessment of gene therapy procedures; in diagnosis
CC (when expressed under control of a promoter induced by a particular
CC analyte); assessment of mutagenicity of compounds; and for drug screening
CC (where expression is controlled by the promoter of a target gene),
CC particularly for antiviral or antiparasitic agents
XX
SQ Sequence 3699 BP; 942 A; 884 C; 898 G; 975 T; 0 U; 0 Other;

```
  Query Match             71.6%;  Score 2620.6;  DB 2;  Length 3699;
  Best Local Similarity   83.3%;  Pred. No. 0;
  Matches 3078;  Conservative    0;  Mismatches  579;  Indels   40;  Gaps    7;

Qy          3 AAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTCAT 62
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3 AAATTGTAAACGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTCAT 62

Qy         63 TTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGAGA 122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 TTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGAGA 122

Qy        123 TAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTCCA 182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 TAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTCCA 182

Qy        183 ACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCCT 242
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 ACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCCT 242

Qy        243 AATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAGCC 302
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 AATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAGCC 302

Qy        303 CCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAAAG 362
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 CCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAAAG 362

Qy        363 CGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCACCA 422
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 CGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCACCA 422

Qy        423 CACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTC---CCATTCGCCATTCAGGCTGC 479
              |||||||||||||||||||||||||||||||||||||    |||||||||||||||||||
Db        423 CACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCGCGCCATTCGCCATTCAGGCTGC 482

Qy        480 GCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAG 539
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        483 GCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAG 542

Qy        540 GGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTT 599
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        543 GGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTT 602

Qy        600 GTAAAACGACGGCCAGTGAGCGCGCCTCGTTCATTCACGTTTTTGAACCCGTGGAGGACG 659
              |||||||||||||||||||     |    | ||||  |   |        ||     |
Db        603 GTAAAACGACGGCCAGTGAATTGTAAT--ACGACTCACTATAGGGCGAATTGGGTACCGG 660

Qy        660 GGCAGACTCGCGGTGCAAATGTGTTTTACAGCGTGATGGAGCAGATGAAGATGCTCGACA 719
              | |  |     | | | || |     ||| |||||   |  ||    ||  | |
Db        661 GCCCCCCCTCGAGGTCGACGGTATCGATAAGCTTGATGATCCTTATTTGTATAGTTCATC 720

Qy        720 CGCTGCA---GAACACGCAGCTAGATTAACCCTAGAAAGATAATCATATTGTGACGTACG 776
              |   ||   | | | ||||    | ||     ||||    |    ||    |
Db        721 CATGCCATGTGTAATCCCAGCAGCTGTTACAAACTCAAGAAGGACCATGTGGTCTCTCTT 780

Qy        777 TTAAAGATAATCATGCGTAAAATTGACGCATGGGATCTGTAATACGACTCACTATAGGGC 836
              ||       ||| | || ||          || |  | |  | | |  | |  ||
Db        781 TTCGTTGGGATCTTTCGAAAGGGCAGATTGTGTGGACAGGTA-ATGGTTGTCTGGTAAAA 839
```

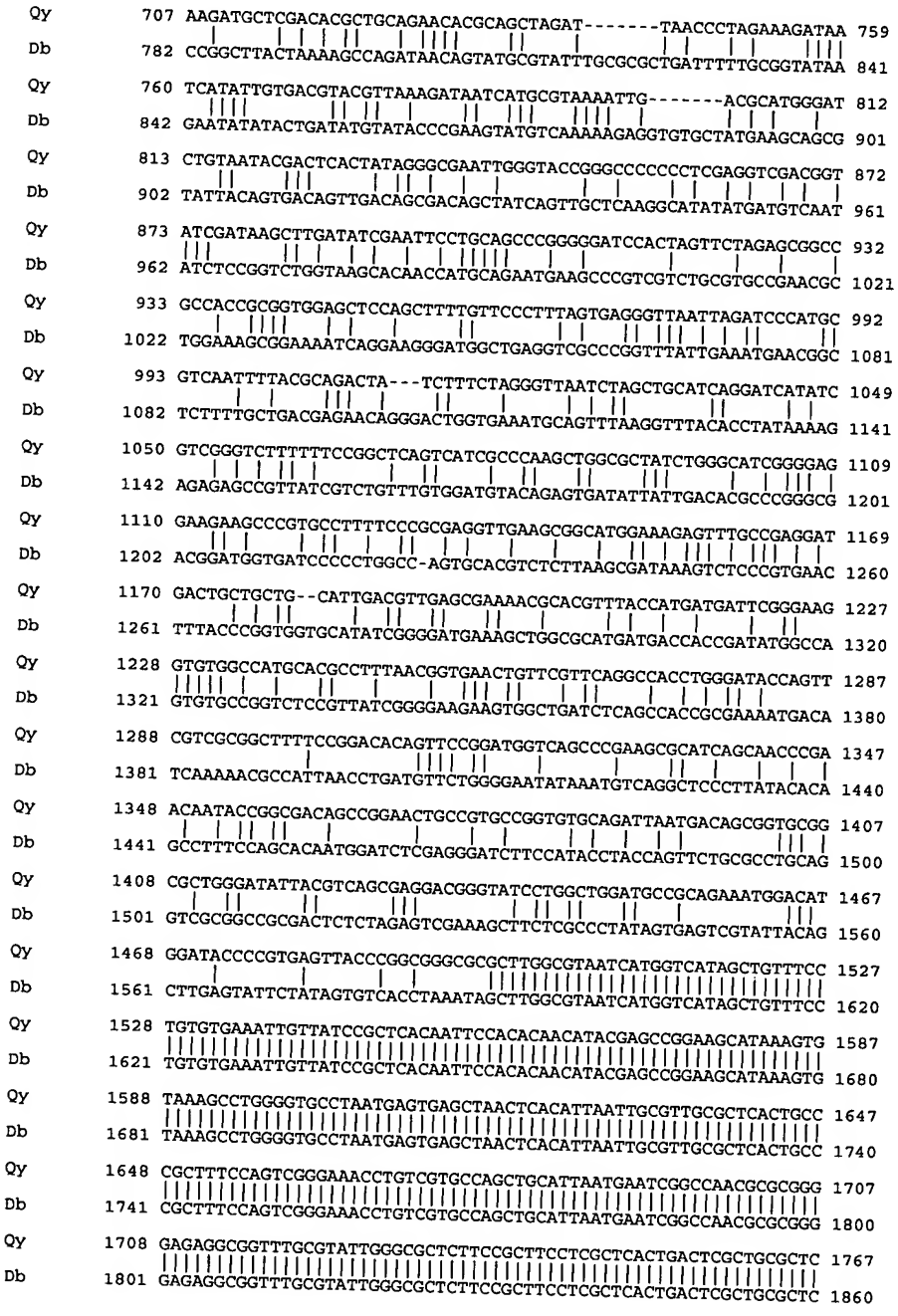


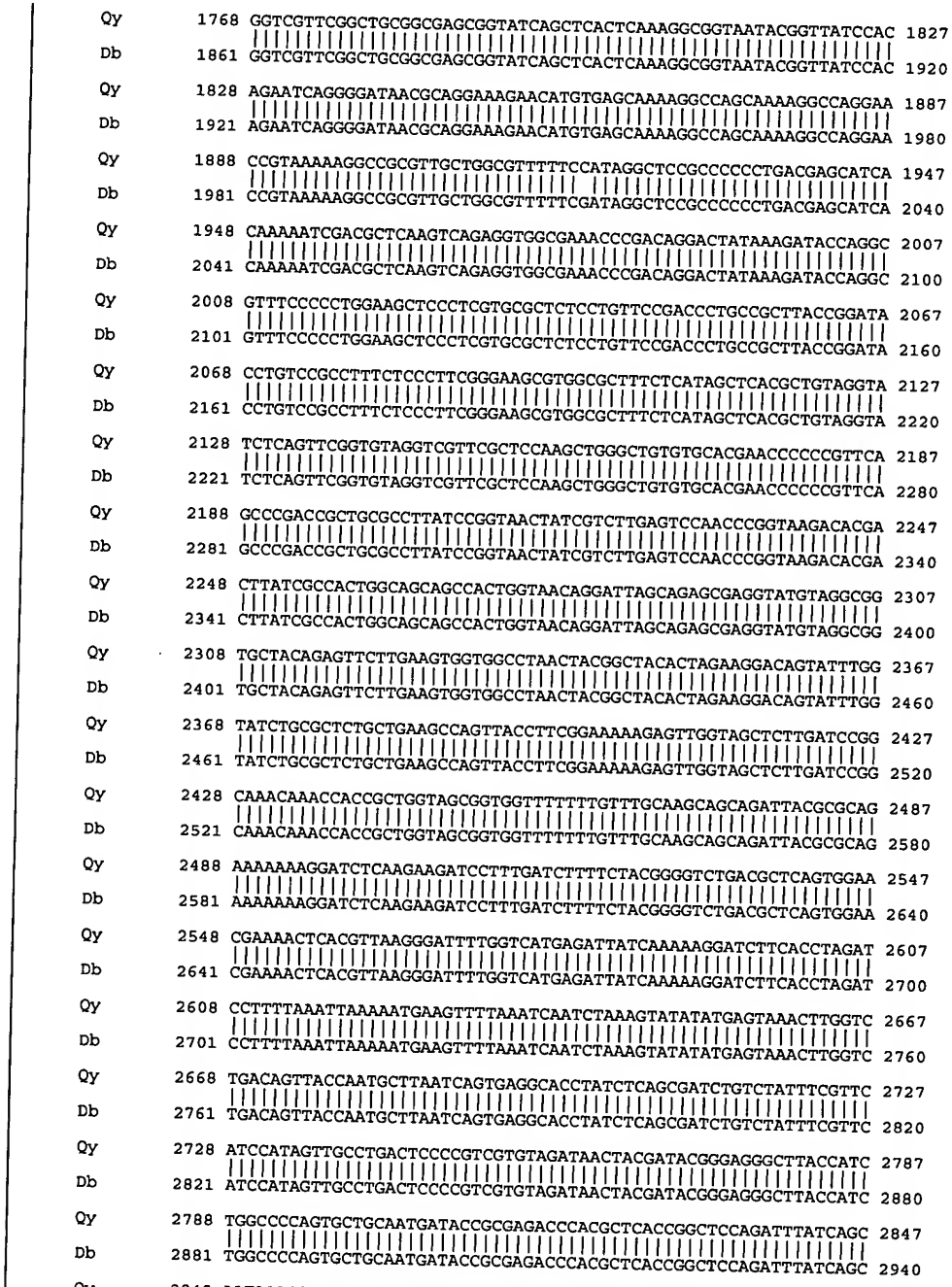
```
Db       3000 TTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATG 3059

Qy       3026 GCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGC 3085

Db       3060 GCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGC 3119

Qy       3086 AAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTG 3145

Db       3120 AAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTG 3179

Qy       3146 TTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGA 3205

Db       3180 TTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGA 3239

Qy       3206 TGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGA 3265

Db       3240 TGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGA 3299

Qy       3266 CCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTA 3325

Db       3300 CCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTA 3359

Qy       3326 AAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTG 3385

Db       3360 AAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTG 3419

Qy       3386 TTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACT 3445

Db       3420 TTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACT 3479

Qy       3446 TTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATA 3505

Db       3480 TTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATA 3539

Qy       3506 AGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATT 3565

Db       3540 AGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATT 3599

Qy       3566 TATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAA 3625

Db       3600 TATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAA 3659

Qy       3626 ATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCAC 3662

Db       3660 ATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCAC 3696
```

RESULT 14
AAD27066
ID AAD27066 standard; DNA; 4001 BP.
XX
AC AAD27066;
XX
DT 09-APR-2002 (first entry)
XX
DE Plasmid pGN59A DNA.
XX
KW Vector construct; RNA inhibition; RNAi; gene expression control;
KW pGN49A plasmid; ds.
XX
OS Unidentified.
XX
PN WO200188121-A1.
XX
PD 22-NOV-2001.
XX
PF 18-MAY-2001; 2001WO-IB001068.
XX
PR 19-MAY-2000; 2000GB-00012233.
XX
PA (DEVG-) DEVGEN NV.
XX
PI Plaetinck G, Renard J, Bogaert T;
XX
DR WPI; 2002-121984/16.
XX
PT A new DNA vector construct containing opposable promoter and terminator
PT sequences flanking a cloning site are useful for the expression of double
PT stranded RNA useful for inhibition of RNA in gene expression control.
XX
PS Claim 24; Fig 15; 75pp; English.
XX
CC The present invention relates to improved vector constructs comprising
CC two promoters in opposite orientation to each other, an inter-promoter
CC region downstream of the 3' end of both promoters, a cloning site in the
CC inter-promoter region and a transcription terminator downstream of the 3'
CC end of the first promoter and the cloning site and operably linked to the
CC first promoter. The constructs of the invention and the bacteria
CC harbouring the constructs are used to produce double stranded RNA for RNA
CC inhibition (RNAi) and can be used as a tool for controlling gene
CC expression. The present sequence is pGN59A plasmid DNA
XX
SQ Sequence 4001 BP; 1034 A; 967 C; 1030 G; 970 T; 0 U; 0 Other;

Query Match 71.3%; Score 2609.4; DB 6; Length 4001;
Best Local Similarity 83.1%; Pred. No. 0;
Matches 3069; Conservative 0; Mismatches 591; Indels 35; Gaps 7;

```
Qy          1 CTAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTC 60

Db         63 CGAAATTGTAAACGTTAATATTTTGTTAAAATTCGCGTTAAATATTTGTTAAATCAGCTC 122

Qy         61 ATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGA 120

Db        123 ATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGA 182

Qy        121 GATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTC 180

Db        183 GATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTC 242

Qy        181 CAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACC 240

Db        243 CAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACC 302

Qy        241 CTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAG 300

Db        303 CAAATCAAGTTTTTTGCGGTCGAGGTGCCGTAAAGCTCTAAATCGGAACCCTAAAGGGAG 362

Qy        301 CCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAA 360

Db        363 CCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAA 422

Qy        361 AGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCAC 420

Db        423 AGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCAC 482

Qy        421 CACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCCCATTCGCCATTCAGGCTGCG 480

Db        483 CACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGT-CCATTCGCCATTCAGGCTGCG 541

Qy        481 CAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGG 540

Db        542 CAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGG 601

Qy        541 GGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTG 600

Db        602 GGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTG 661

Qy        601 TAAAACGACGGCCAGTGAGCGCGCCT--------------CGTTCATTCACGTTTTTGAA 646

Db        662 TAAAACGACGGCCAGTGAATTGTAATACGACTCACTATAGGGCGAATTCGAGCTCGGTAC 721

Qy        647 CCCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGCGTGATGGAGCAGATG 706

Db        722 CCGGGGATCCTCTAGAGATCCCTCGACCTCGAGATCCATTGTGCTGGAAAGGATCTGGAT 781
```



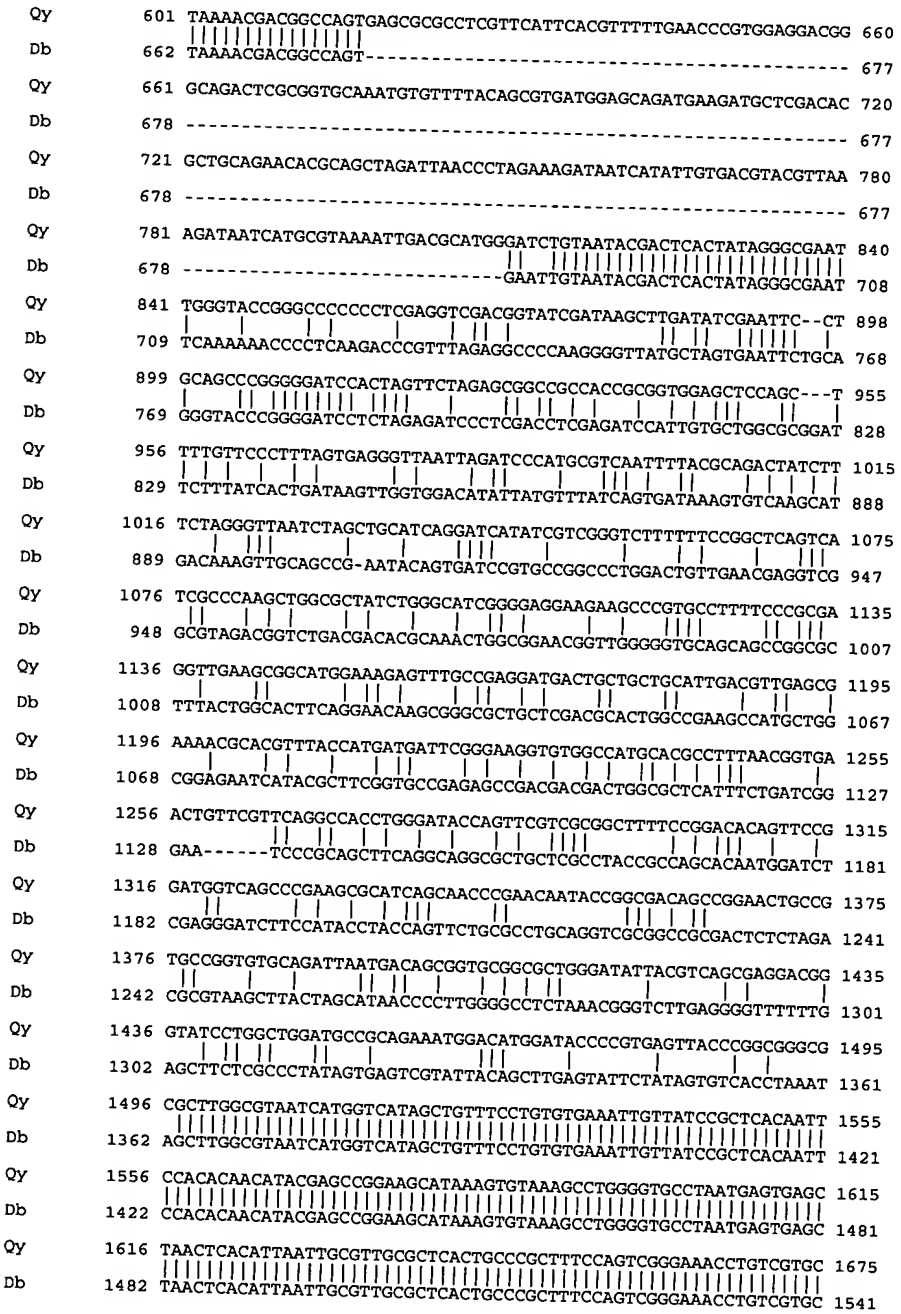


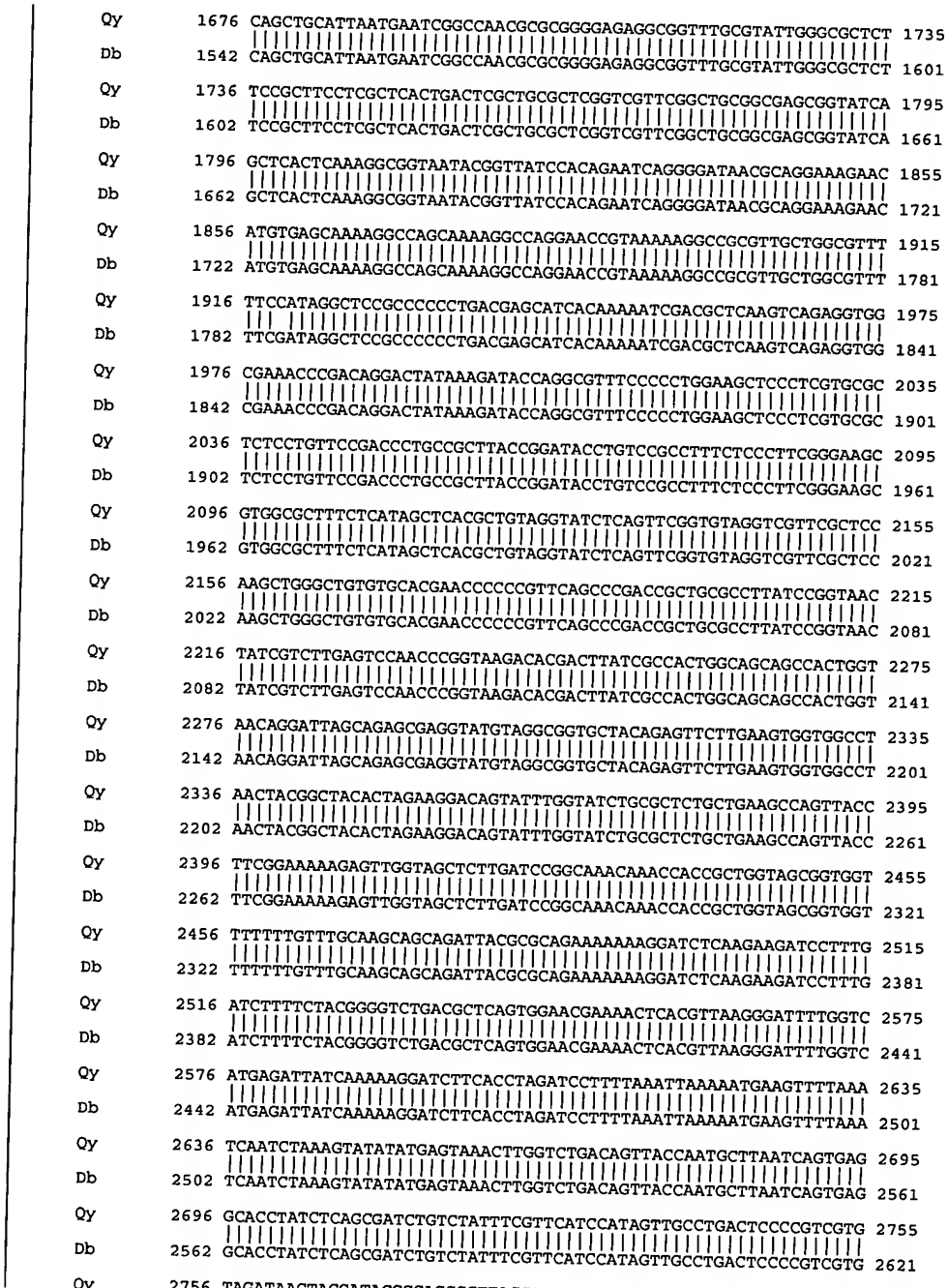
```
Db       2941 AATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTC 3000

Qy       2908 CATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTT 2967
Db       3001 CATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTT 3060

Qy       2968 GCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGC 3027
Db       3061 GCGCAACGTTGTTGGCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGC 3120

Qy       3028 TTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAA 3087
Db       3121 TTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAA 3180

Qy       3088 AAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTT 3147
Db       3181 AAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTT 3240

Qy       3148 ATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATG 3207
Db       3241 ATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATG 3300

Qy       3208 CTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACC 3267
Db       3301 CTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATACCGCGCCCGGCGACC 3360

Qy       3268 GAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAA 3327
Db       3361 GAGTTGCTCTTGCCCGGCGTCAATACGGGATAATAGTGTATGACATAGCAGAACTTTAAA 3420

Qy       3328 AGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTT 3387
Db       3421 AGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTT 3480

Qy       3388 GAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTT 3447
Db       3481 GAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTT 3540

Qy       3448 CACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAG 3507
Db       3541 CACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAG 3600

Qy       3508 GGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTA 3567
Db       3601 GGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTA 3660

Qy       3568 TCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAAT 3627
Db       3661 TCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAAT 3720

Qy       3628 AGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCAC 3662
Db       3721 AGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCAC 3755
```

RESULT 15
AAD27062
ID AAD27062 standard; DNA; 3774 BP.
XX
AC AAD27062;
XX
DT 09-APR-2002 (first entry)
XX
DE Plasmid pGN29 DNA.
XX
KW Vector construct; RNA inhibition; RNAi; gene expression control;
KW pGN29 plasmid; ds.
XX
OS Unidentified.
XX
PN WO200188121-A1.
XX
PD 22-NOV-2001.
XX
PF 18-MAY-2001; 2001WO-IB001068.
XX
PR 19-MAY-2000; 2000GB-00012233.
XX
PA (DEVG-) DEVGEN NV.
XX
PI Plaetinck G, Renard J, Bogaert T;
XX
DR WPI; 2002-121984/16.
XX
PT A new DNA vector construct containing opposable promoter and terminator
PT sequences flanking a cloning site are useful for the expression of double
PT stranded RNA useful for inhibition of RNA in gene expression control.
XX
PS Claim 24; Fig 11; 75pp; English.
XX
CC The present invention relates to improved vector constructs comprising
CC two promoters in opposite orientation to each other, an inter-promoter
CC region downstream of the 3' end of both promoters, a cloning site in the
CC inter-promoter region and a transcription terminator downstream of the 3'
CC end of the first promoter and the cloning site and operably linked to the
CC first promoter. The constructs of the invention and the bacteria
CC harbouring the constructs are used to produce double stranded RNA for RNA
CC inhibition (RNAi) and can be used as a tool for controlling gene
CC expression. The present sequence is pGN29 plasmid DNA
XX
SQ Sequence 3774 BP; 948 A; 940 C; 983 G; 903 T; 0 U; 0 Other;

Query Match 67.8%; Score 2484.6; DB 6; Length 3774;
Best Local Similarity 82.3%; Pred. No. 0;
Matches 3017; Conservative 0; Mismatches 444; Indels 206; Gaps 6;

```
Qy          1 CTAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTC 60
Db         63 CGAAATTGTAAACGTTAATATTTTGTTAAAATTCGCGTTAAATATTTGTTAAATCAGCTC 122

Qy         61 ATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGA 120
Db        123 ATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGA 182

Qy        121 GATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTC 180
Db        183 GATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTC 242

Qy        181 CAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACC 240
Db        243 CAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACC 302

Qy        241 CTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAG 300
Db        303 CAAATCAAGTTTTTTGCGGTCGAGGTGCCGTAAAGCTCTAAATCGGAACCCTAAAGGGAG 362

Qy        301 CCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAA 360
Db        363 CCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAA 422

Qy        361 AGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCAC 420
Db        423 AGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCAC 482

Qy        421 CACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCCCATTCGCCATTCAGGCTGCG 480
Db        483 CACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGT-CCATTCGCCATTCAGGCTGCG 541

Qy        481 CAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGG 540
Db        542 CAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGG 601

Qy        541 GGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTG 600
Db        602 GGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTG 661
```





```
Db       2622 TAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGA 2681

Qy       2816 GACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAG 2875

Db       2682 GACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAG 2741

Qy       2876 CGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAA 2935

Db       2742 CGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAA 2801

Qy       2936 GCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGC 2995

Db       2802 GCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGGCATTGCTACAGGC 2861

Qy       2996 ATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCA 3055

Db       2862 ATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCA 2921

Qy       3056 AGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCG 3115

Db       2922 AGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCG 2981

Qy       3116 ATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCAT 3175

Db       2982 ATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCAT 3041

Qy       3176 AATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACC 3235

Db       3042 AATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACC 3101

Qy       3236 AAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGG 3295

Db       3102 AAGTCATTCTGAGAATACCGCGCCCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGG 3161

Qy       3296 GATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCG 3355

Db       3162 GATAATAGTGTATGACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCG 3221

Qy       3356 GGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGT 3415

Db       3222 GGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGT 3281

Qy       3416 GCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACA 3475

Db       3282 GCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACA 3341

Qy       3476 GGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATA 3535

Db       3342 GGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATA 3401

Qy       3536 CTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATAC 3595

Db       3402 CTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATAC 3461

Qy       3596 ATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAA 3655

Db       3462 ATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAA 3521

Qy       3656 GTGCCAC 3662

Db       3522 GTGCCAC 3528


Search completed: January 26, 2005, 03:15:10
Job time : 1771.03 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: January 26, 2005, 08:26:03 ; Search time 1959.66 Seconds
(without alignments)
10737.291 Million cell updates/sec

Title: US-10-826-523-41
Perfect score: 3662
Sequence: 1 ctaaattgtaagcgttaata.....atttccccgaaaagtgccac 3662

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4300275 seqs, 2872944193 residues

Total number of hits satisfying chosen parameters: 8600550

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
20: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
21: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
	1	3662	100.0	3662	13	US-10-001-189-41	Sequence 41, Appl
c	2	2665.4	72.8	3637	15	US-10-211-079-18	Sequence 18, Appl
c	3	2665.4	72.8	3637	15	US-10-356-708-32	Sequence 32, Appl
c	4	2665.4	72.8	3637	17	US-10-280-913A-32	Sequence 32, Appl
c	5	2665.4	72.8	3637	17	US-10-684-134-32	Sequence 32, Appl
c	6	2665.4	72.8	3637	17	US-10-637-758-32	Sequence 32, Appl
c	7	2664.8	72.8	3637	13	US-10-066-390-4	Sequence 4, Appli
c	8	2664.8	72.8	3637	13	US-10-206-030-4	Sequence 4, Appli
c	9	2664.8	72.8	3637	15	US-10-211-079-4	Sequence 4, Appli
c	10	2664.8	72.8	3637	15	US-10-356-708-17	Sequence 17, Appl
c	11	2664.8	72.8	3637	15	US-10-205-772-4	Sequence 4, Appli
c	12	2664.8	72.8	3637	17	US-10-280-913A-17	Sequence 17, Appl
c	13	2664.8	72.8	3637	17	US-10-684-134-17	Sequence 17, Appl
c	14	2664.8	72.8	3637	17	US-10-637-758-17	Sequence 17, Appl
c	15	2656.4	72.5	3637	13	US-10-066-390-3	Sequence 3, Appli
c	16	2656.4	72.5	3637	13	US-10-206-030-3	Sequence 3, Appli
c	17	2656.4	72.5	3637	15	US-10-211-079-3	Sequence 3, Appli
c	18	2656.4	72.5	3637	15	US-10-356-708-31	Sequence 31, Appl
c	19	2656.4	72.5	3637	15	US-10-205-772-3	Sequence 3, Appli
c	20	2656.4	72.5	3637	17	US-10-280-913A-31	Sequence 31, Appl
c	21	2656.4	72.5	3637	17	US-10-684-134-31	Sequence 31, Appl
c	22	2656.4	72.5	3637	17	US-10-637-758-31	Sequence 31, Appl
	23	2609.4	71.3	4001	11	US-09-860-763-13	Sequence 13, Appl
	24	2484.6	67.8	3774	11	US-09-860-763-9	Sequence 9, Appli
	25	2469.4	67.4	5277	15	US-10-412-872-16	Sequence 16, Appl
	26	2464.6	67.3	3908	16	US-10-014-099F-70	Sequence 70, Appl
	27	2451.8	67.0	3928	15	US-10-128-590-45	Sequence 45, Appl
	28	2451.8	67.0	3928	15	US-10-128-587A-45	Sequence 45, Appl
	29	2451.8	67.0	3928	15	US-10-128-578B-45	Sequence 45, Appl
	30	2446.8	66.8	3715	11	US-09-860-763-11	Sequence 11, Appl
	31	2446.8	66.8	3927	16	US-10-014-099F-71	Sequence 71, Appl
	32	2279.2	62.2	4119	9	US-09-993-170-1	Sequence 1, Appli
	33	2271.2	62.0	4059	17	US-10-655-702-1	Sequence 1, Appli
	34	2230.6	60.9	5250	18	US-10-474-070-51	Sequence 51, Appl
	35	2216.2	60.5	4205	10	US-09-883-848A-48	Sequence 48, Appl
	36	2213	60.4	4205	10	US-09-883-848A-35	Sequence 35, Appl
c	37	2172.8	59.3	7803	16	US-10-014-099F-82	Sequence 82, Appl
c	38	2172.8	59.3	8167	16	US-10-014-099F-83	Sequence 83, Appl
	39	2172.6	59.3	10600	15	US-10-356-708-1	Sequence 1, Appli
	40	2172.6	59.3	10600	17	US-10-280-913A-1	Sequence 1, Appli
	41	2172.6	59.3	10600	17	US-10-684-134-1	Sequence 1, Appli
	42	2172.6	59.3	10600	17	US-10-637-758-1	Sequence 1, Appli
	43	2172.6	59.3	10624	15	US-10-356-708-2	Sequence 2, Appli
	44	2172.6	59.3	10624	17	US-10-280-913A-2	Sequence 2, Appli
	45	2172.6	59.3	10624	17	US-10-684-134-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-10-001-189-41
; Sequence 41, Application US/10001189
; Publication No. US20020173634A1
; GENERAL INFORMATION:
; APPLICANT: FRASER JR., MALCOLM J.
; APPLICANT: LI, XU
; APPLICANT: BEAM, TERESA
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPOSITION USING
; TITLE OF INVENTION: MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION
; TITLE OF INVENTION: VECTOR PIGGYBAC
; FILE REFERENCE: 835910-92098
; CURRENT APPLICATION NUMBER: US/10/001,189
; CURRENT FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: 60/244,984
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/244,677
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 41
; LENGTH: 3662
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: pXL-Bac
; OTHER INFORMATION: sequence
US-10-001-189-41

Query Match 100.0%; Score 3662; DB 13; Length 3662;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3662; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

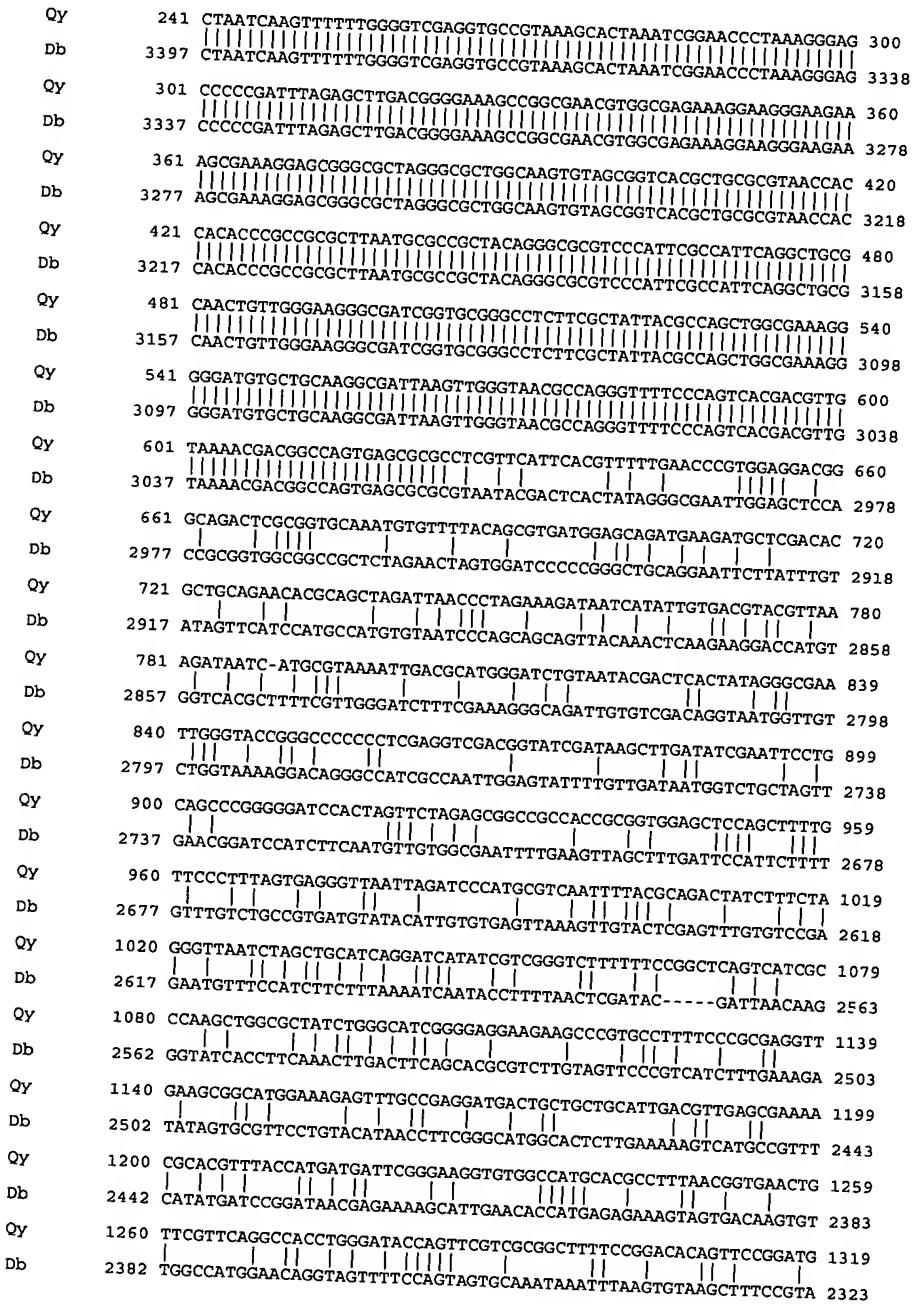
Qy 1 CTAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTC 60
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

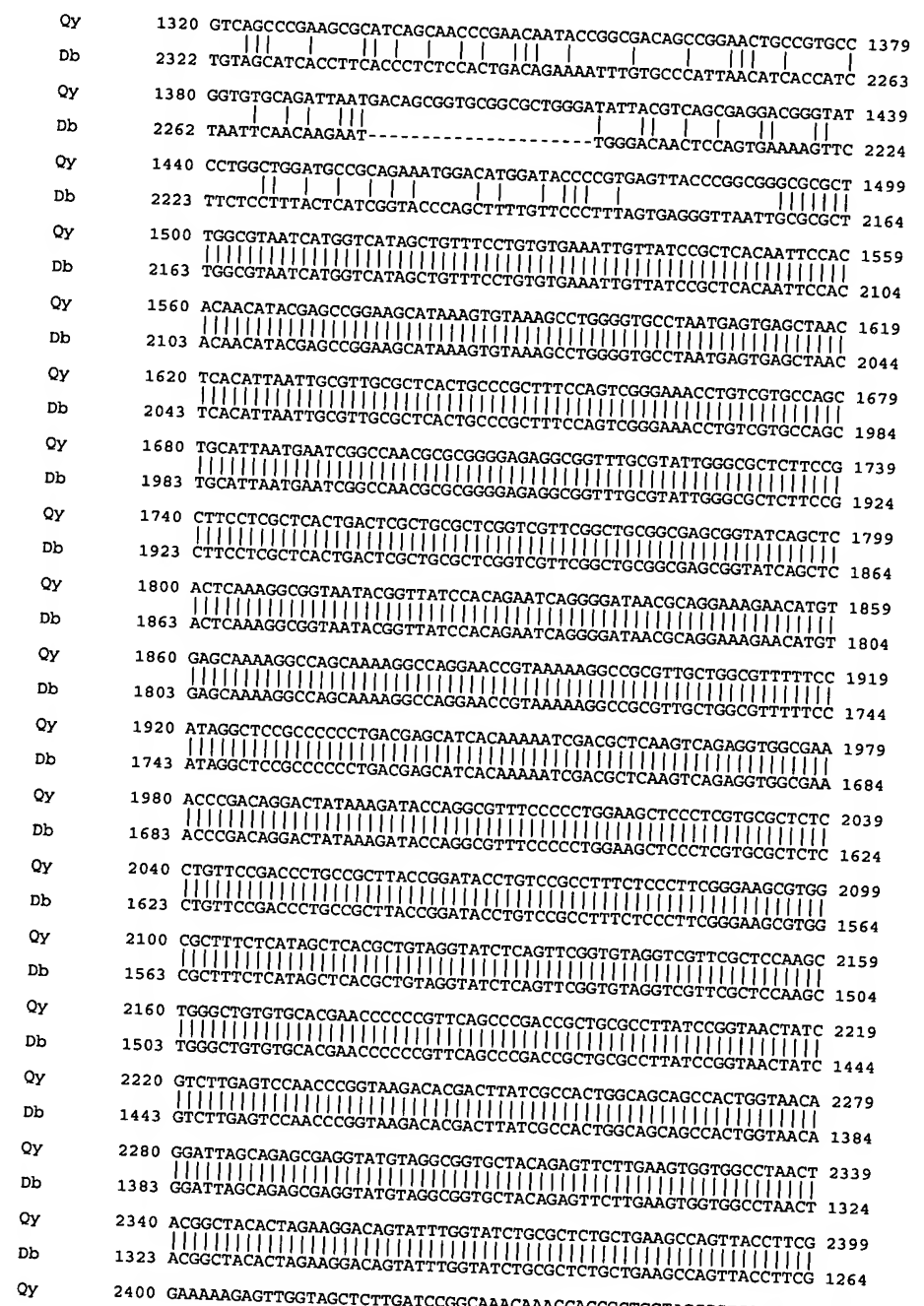
```
Qy       2221 TCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAG 2280
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 TCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAG 2280

Qy       2281 GATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTA 2340
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 GATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTA 2340

Qy       2341 CGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGG 2400
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2341 CGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGG 2400

Qy       2401 AAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTT 2460
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2401 AAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTT 2460

Qy       2461 TGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTT 2520
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2461 TGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTT 2520

Qy       2521 TTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAG 2580
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 TTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAG 2580

Qy       2581 ATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAAT 2640
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2581 ATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAAT 2640

Qy       2641 CTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACC 2700
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 CTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACC 2700

Qy       2701 TATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGAT 2760
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 TATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGAT 2760

Qy       2761 AACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCC 2820
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 AACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCC 2820

Qy       2821 ACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAG 2880
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 ACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAG 2880

Qy       2881 AAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAG 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2881 AAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAG 2940

Qy       2941 AGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGT 3000
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2941 AGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGT 3000

Qy       3001 GGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCG 3060
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3001 GGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCG 3060

Qy       3061 AGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGT 3120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3061 AGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGT 3120

Qy       3121 TGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTC 3180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3121 TGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTC 3180

Qy       3181 TCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTC 3240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3181 TCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTC 3240

Qy       3241 ATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAA 3300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3241 ATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAA 3300

Qy       3301 TACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCG 3360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3301 TACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCG 3360

Qy       3361 AAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACC 3420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3361 AAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACC 3420

Qy       3421 CAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAG 3480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3421 CAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAG 3480

Qy       3481 GCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTT 3540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3481 GCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTT 3540

Qy       3541 CCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATT 3600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3541 CCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATT 3600

Qy       3601 TGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCC 3660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3601 TGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCC 3660

Qy       3661 AC 3662
              ||
Db       3661 AC 3662
```

RESULT 2
US-10-211-079-18/c
; Sequence 18, Application US/10211079
; Publication No. US20030148315A1
; GENERAL INFORMATION:
; APPLICANT: Padgett, Hal S.
; APPLICANT: Vaewhongs, Andrew A.
; APPLICANT: Vojdani, Fakhrieh S.
; APPLICANT: Smith, Mark L.
; TITLE OF INVENTION: Nucleic Acid Molecules Encoding CEL I
; TITLE OF INVENTION: Endonuclease and Methods of Use Thereof
; FILE REFERENCE: P-LG 5381
; CURRENT APPLICATION NUMBER: US/10/211,079
; CURRENT FILING DATE: 2002-08-01
; PRIOR APPLICATION NUMBER: US 10/098,155
; PRIOR FILING DATE: 2002-03-14
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 3637
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic construct
US-10-211-079-18

Query Match 72.8%; Score 2665.4; DB 15; Length 3637;
Best Local Similarity 83.7%; Pred. No. 0;
Matches 3065; Conservative 0; Mismatches 571; Indels 27; Gaps 3;

```
Qy          1 CTAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3637 CTAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTC 3578

Qy         61 ATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3577 ATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGA 3518

Qy        121 GATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3517 GATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTC 3458

Qy        181 CAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3457 CAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACC 3398
```





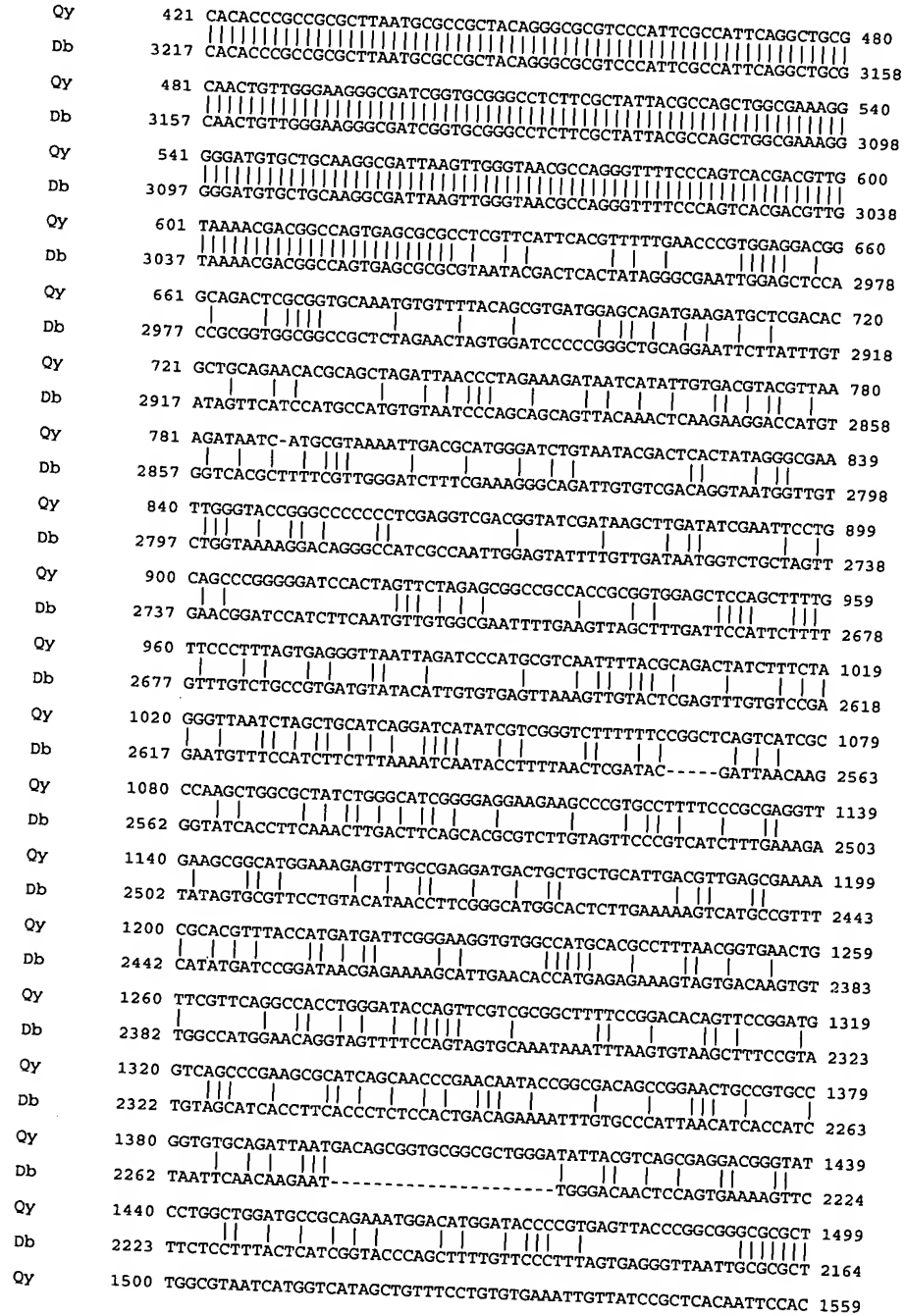
```
Db       1263 GAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTT 1204

Qy       2460 TTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCT 2519
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1203 TTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCT 1144

Qy       2520 TTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGA 2579
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1143 TTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGA 1084

Qy       2580 GATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAA 2639
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1083 GATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAA 1024

Qy       2640 TCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCAC 2699
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1023 TCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCAC 964

Qy       2700 CTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGA 2759
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        963 CTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGA 904

Qy       2760 TAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACC 2819
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        903 TAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACC 844

Qy       2820 CACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCA 2879
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        843 CACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCA 784

Qy       2880 GAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTA 2939
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        783 GAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTA 724

Qy       2940 GAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCG 2999
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        723 GAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCG 664

Qy       3000 TGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGC 3059
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        663 TGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGC 604

Qy       3060 GAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCG 3119
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        603 GAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCG 544

Qy       3120 TTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATT 3179
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        543 TTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATT 484

Qy       3180 CTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGT 3239
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        483 CTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGT 424

Qy       3240 CATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATA 3299
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        423 CATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATA 364

Qy       3300 ATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGC 3359
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 ATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGC 304

Qy       3360 GAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCAC 3419
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 GAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCAC 244

Qy       3420 CCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAA 3479
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 CCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAA 184

Qy       3480 GGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCT 3539
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        183 GGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCT 124

Qy       3540 TCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATAT 3599
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 TCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATAT 64

Qy       3600 TTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGC 3659
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 TTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGC 4

Qy       3660 CAC 3662
              |||
Db          3 CAC 1
```

RESULT 3
US-10-356-708-32/c
; Sequence 32, Application US/10356708
; Publication No. US20030157682A1
; GENERAL INFORMATION:
; APPLICANT: Large Scale Biology Corporation
; TITLE OF INVENTION: MISMATCH ENDONUCLEASES AND METHODS OF USE THEREOF
; FILE REFERENCE: P-LG 10100
; CURRENT APPLICATION NUMBER: US/10/356,708
; CURRENT FILING DATE: 2003-01-31
; PRIOR APPLICATION NUMBER: 60/353,722
; PRIOR FILING DATE: 2000-02-01
; PRIOR APPLICATION NUMBER: 10/098,155
; PRIOR FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 10/211,079
; PRIOR FILING DATE: 2002-08-01
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 32
; LENGTH: 3637
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Encodes Cycle 3 BFP gene
US-10-356-708-32

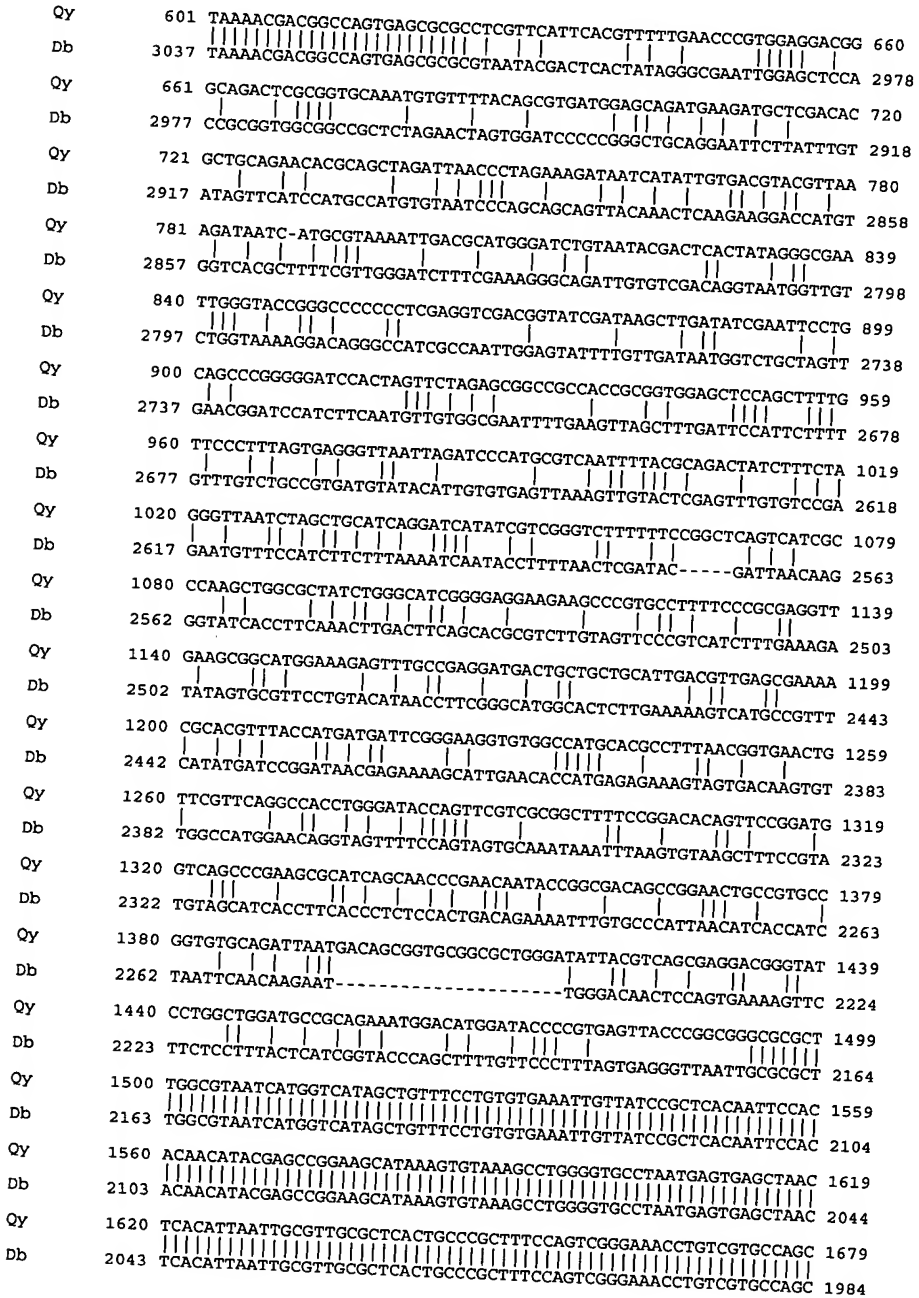
Query Match 72.8%; Score 2665.4; DB 15; Length 3637;
Best Local Similarity 83.7%; Pred. No. 0;
Matches 3065; Conservative 0; Mismatches 571; Indels 27; Gaps 3;

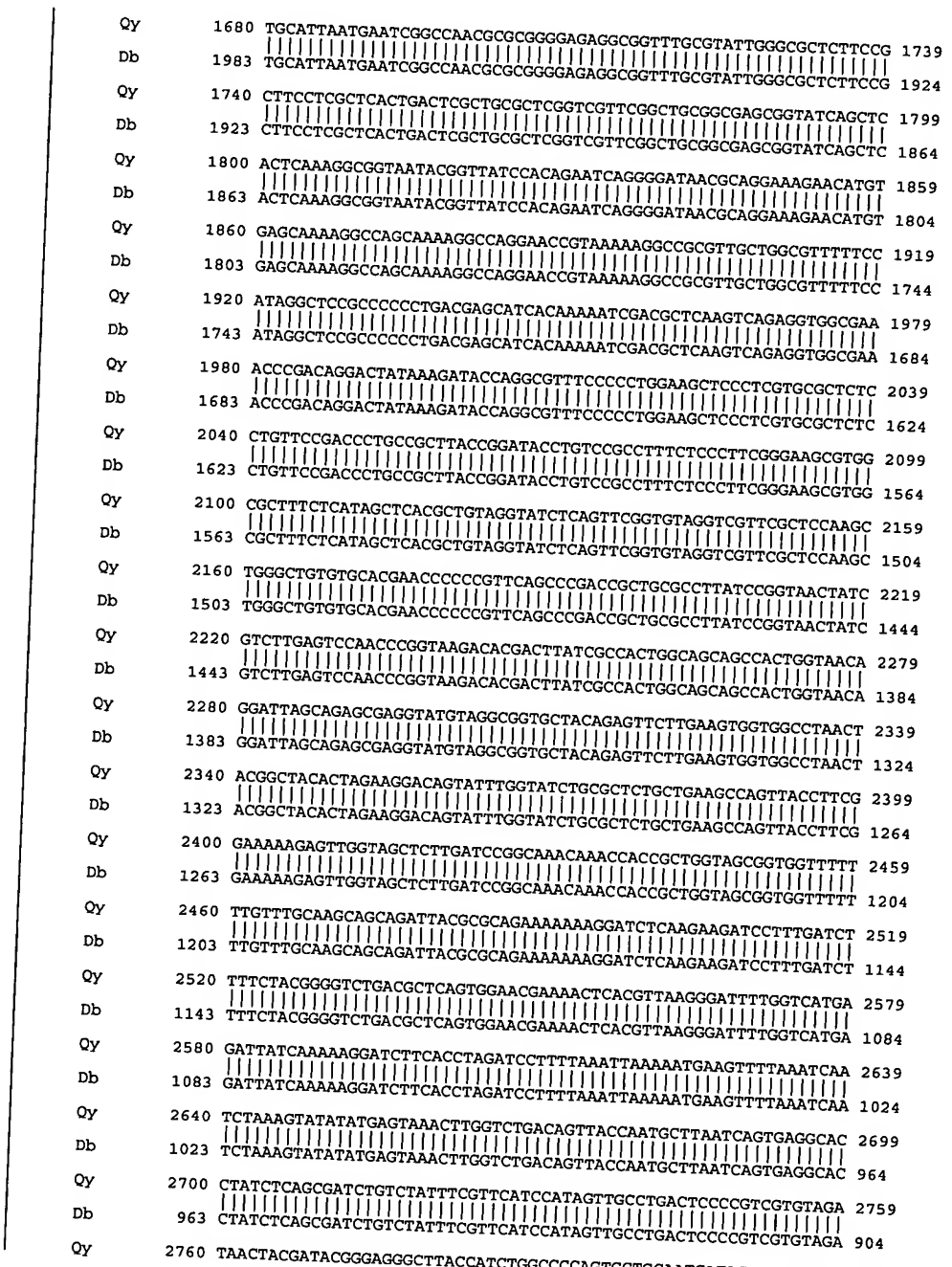
```
Qy          1 CTAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3637 CTAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTC 3578

Qy         61 ATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3577 ATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGA 3518

Qy        121 GATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3517 GATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTC 3458

Qy        181 CAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3457 CAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACC 3398

Qy        241 CTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3397 CTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAG 3338

Qy        301 CCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3337 CCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAA 3278

Qy        361 AGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCAC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3277 AGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCAC 3218
```


Qy 421 CACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCCCATTCGCCATTCAGGCTGCG 480
Db 3217 CACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCCCATTCGCCATTCAGGCTGCG 3158
Qy 481 CAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGG 540
Db 3157 CAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGG 3098
Qy 541 GGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTG 600
Db 3097 GGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTG 3038
Qy 601 TAAAACGACGGCCAGTGAGCGCGCCTCGTTCATTCACGTTTTTGAACCCGTGGAGGACGG 660
Db 3037 TAAAACGACGGCCAGTGAGCGCGCGTAATACGACTCACTATAGGGCGAATTGGAGCTCCA 2978
Qy 661 GCAGACTCGCGGTGCAAATGTGTTTTACAGCGTGATGGAGCAGATGAAGATGCTCGACAC 720
Db 2977 CCGCGGTGGCGGCCGCTCTAGAACTAGTGGATCCCCCGGGCTGCAGGAATTCTTATTTGT 2918
Qy 721 GCTGCAGAACACGCAGCTAGATTAACCCTAGAAAGATAATCATATTGTGACGTACGTTAA 780
Db 2917 ATAGTTCATCCATGCCATGTGTAATCCCAGCAGCAGTTACAAACTCAAGAAGGACCATGT 2858
Qy 781 AGATAATC-ATGCGTAAAATTGACGCATGGGATCTGTAATACGACTCACTATAGGGCGAA 839
Db 2857 GGTCACGCTTTTCGTTGGGATCTTTCGAAAGGGCAGATTGTGTCGACAGGTAATGGTTGT 2798
Qy 840 TTGGGTACCGGGCCCCCCCTCGAGGTCGACGGTATCGATAAGCTTGATATCGAATTCCTG 899
Db 2797 CTGGTAAAAGGACAGGGCCATCGCCAATTGGAGTATTTTGTTGATAATGGTCTGCTAGTT 2738
Qy 900 CAGCCCGGGGGATCCACTAGTTCTAGAGCGGCCGCCACCGCGGTGGAGCTCCAGCTTTTG 959
Db 2737 GAACGGATCCATCTTCAATGTTGTGGCGAATTTTGAAGTTAGCTTTGATTCCATTCTTTT 2678
Qy 960 TTCCCTTTAGTGAGGGTTAATTAGATCCCATGCGTCAATTTTACGCAGACTATCTTTCTA 1019
Db 2677 GTTTGTCTGCCGTGATGTATACATTGTGTGAGTTAAAGTTGTACTCGAGTTTGTGTCCGA 2618
Qy 1020 GGGTTAATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGC 1079
Db 2617 GAATGTTTCCATCTTCTTTAAAATCAATACCTTTTAACTCGATAC-----GATTAACAAG 2563
Qy 1080 CCAAGCTGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTT 1139
Db 2562 GGTATCACCTTCAAACTTGACTTCAGCACGCGTCTTGTAGTTCCCGTCATCTTTGAAAGA 2503
Qy 1140 GAAGCGGCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAA 1199
Db 2502 TATAGTGCGTTCCTGTACATAACCTTCGGGCATGGCACTCTTGAAAAAGTCATGCCGTTT 2443
Qy 1200 CGCACGTTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTG 1259
Db 2442 CATATGATCCGGATAACGAGAAAAGCATTGAACACCATGAGAGAAAGTAGTGACAAGTGT 2383
Qy 1260 TTCGTTCAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATG 1319
Db 2382 TGGCCATGGAACAGGTAGTTTTCCAGTAGTGCAAATAAATTTAAGTGTAAGCTTTCCGTA 2323
Qy 1320 GTCAGCCCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCC 1379
Db 2322 TGTAGCATCACCTTCACCCTCTCCACTGACAGAAAATTTGTGCCCATTAACATCACCATC 2263
Qy 1380 GGTGTGCAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTAT 1439
Db 2262 TAATTCAACAAGAAT----------TGGGACAACTCCAGTGAAAAGTTC 2224
Qy 1440 CCTGGCTGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCT 1499
Db 2223 TTCTCCTTTACTCATCGGTACCCAGCTTTTGTTCCCTTTAGTGAGGGTTAATTGCGCGCT 2164
Qy 1500 TGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCAC 1559
Db 2163 TGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCAC 2104
Qy 1560 ACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAAC 1619
Db 2103 ACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAAC 2044
Qy 1620 TCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGC 1679
Db 2043 TCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGC 1984
Qy 1680 TGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCG 1739
Db 1983 TGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCG 1924
Qy 1740 CTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTC 1799
Db 1923 CTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTC 1864
Qy 1800 ACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGT 1859
Db 1863 ACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGT 1804
Qy 1860 GAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCC 1919
Db 1803 GAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCC 1744
Qy 1920 ATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAA 1979
Db 1743 ATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAA 1684
Qy 1980 ACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTC 2039
Db 1683 ACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTC 1624
Qy 2040 CTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGG 2099
Db 1623 CTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGG 1564
Qy 2100 CGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGC 2159
Db 1563 CGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGC 1504
Qy 2160 TGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATC 2219
Db 1503 TGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATC 1444
Qy 2220 GTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACA 2279
Db 1443 GTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACA 1384
Qy 2280 GGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACT 2339
Db 1383 GGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACT 1324
Qy 2340 ACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCG 2399
Db 1323 ACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCG 1264
Qy 2400 GAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTT 2459
Db 1263 GAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTT 1204
Qy 2460 TTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCT 2519
Db 1203 TTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCT 1144
Qy 2520 TTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGA 2579
Db 1143 TTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGA 1084
Qy 2580 GATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAA 2639

```
Db       1083 GATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAA 1024

Qy       2640 TCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCAC 2699
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1023 TCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCAC 964

Qy       2700 CTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGA 2759
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        963 CTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGA 904

Qy       2760 TAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACC 2819
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        903 TAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACC 844

Qy       2820 CACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCA 2879
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        843 CACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCA 784

Qy       2880 GAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTA 2939
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        783 GAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTA 724

Qy       2940 GAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCG 2999
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        723 GAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCG 664

Qy       3000 TGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGC 3059
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        663 TGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGC 604

Qy       3060 GAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCG 3119
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        603 GAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCG 544

Qy       3120 TTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATT 3179
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        543 TTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATT 484

Qy       3180 CTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGT 3239
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        483 CTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGT 424

Qy       3240 CATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATA 3299
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        423 CATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATA 364

Qy       3300 ATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGC 3359
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 ATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGC 304

Qy       3360 GAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCAC 3419
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 GAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCAC 244

Qy       3420 CCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAA 3479
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 CCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAA 184

Qy       3480 GGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCT 3539
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 GGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCT 124

Qy       3540 TCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATAT 3599
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 TCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATAT 64

Qy       3600 TTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGC 3659
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          63 TTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGC 4

Qy       3660 CAC 3662
             |||
Db           3 CAC 1
```

```
RESULT 4
US-10-280-913A-32/c
; Sequence 32, Application US/10280913A
; Publication No. US20040110130A1
; GENERAL INFORMATION:
;  APPLICANT: Large Scale Biology Corporation
;  TITLE OF INVENTION: POPULATION OF POLYNUCLEOTIDE SEQUENCE VARIANTS
;  FILE REFERENCE: LSBC-PLG5392-CIP
;  CURRENT APPLICATION NUMBER: US/10/280,913A
;  CURRENT FILING DATE:  2002-10-25
;  PRIOR APPLICATION NUMBER: 60/402,342
;  PRIOR FILING DATE: 2002-08-08
;  PRIOR APPLICATION NUMBER: 10/066,390
;  PRIOR FILING DATE: 2002-02-01
;  PRIOR APPLICATION NUMBER: 60/268,785
;  PRIOR FILING DATE: 2001-02-14
;  PRIOR APPLICATION NUMBER: 60/266,386
;  PRIOR FILING DATE: 2002-02-02
;  NUMBER OF SEQ ID NOS: 35
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 32
;   LENGTH: 3637
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Encodes Cycle 3 BFP gene
US-10-280-913A-32

  Query Match             72.8%;  Score 2665.4;  DB 17;  Length 3637;
  Best Local Similarity   83.7%;  Pred. No. 0;
  Matches 3065;  Conservative    0;  Mismatches  571;  Indels   27;  Gaps    3;

Qy          1 CTAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3637 CTAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTC 3578

Qy         61 ATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3577 ATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGA 3518

Qy        121 GATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3517 GATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTC 3458

Qy        181 CAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3457 CAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACC 3398

Qy        241 CTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3397 CTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAG 3338

Qy        301 CCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3337 CCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAA 3278

Qy        361 AGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCAC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3277 AGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCAC 3218

Qy        421 CACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCCCATTCGCCATTCAGGCTGCG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3217 CACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCCCATTCGCCATTCAGGCTGCG 3158

Qy        481 CAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3157 CAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGG 3098

Qy        541 GGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3097 GGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTG 3038
```



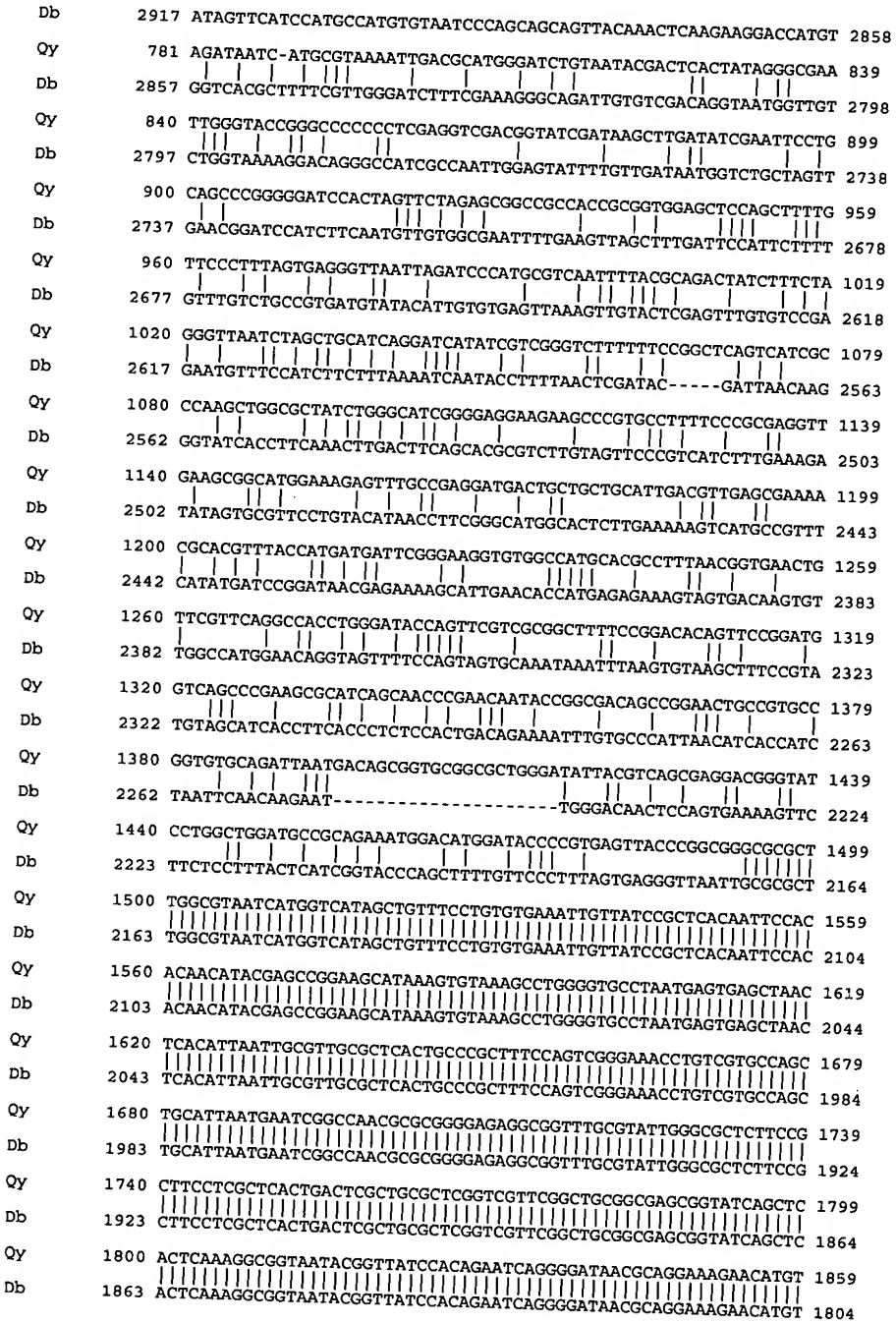


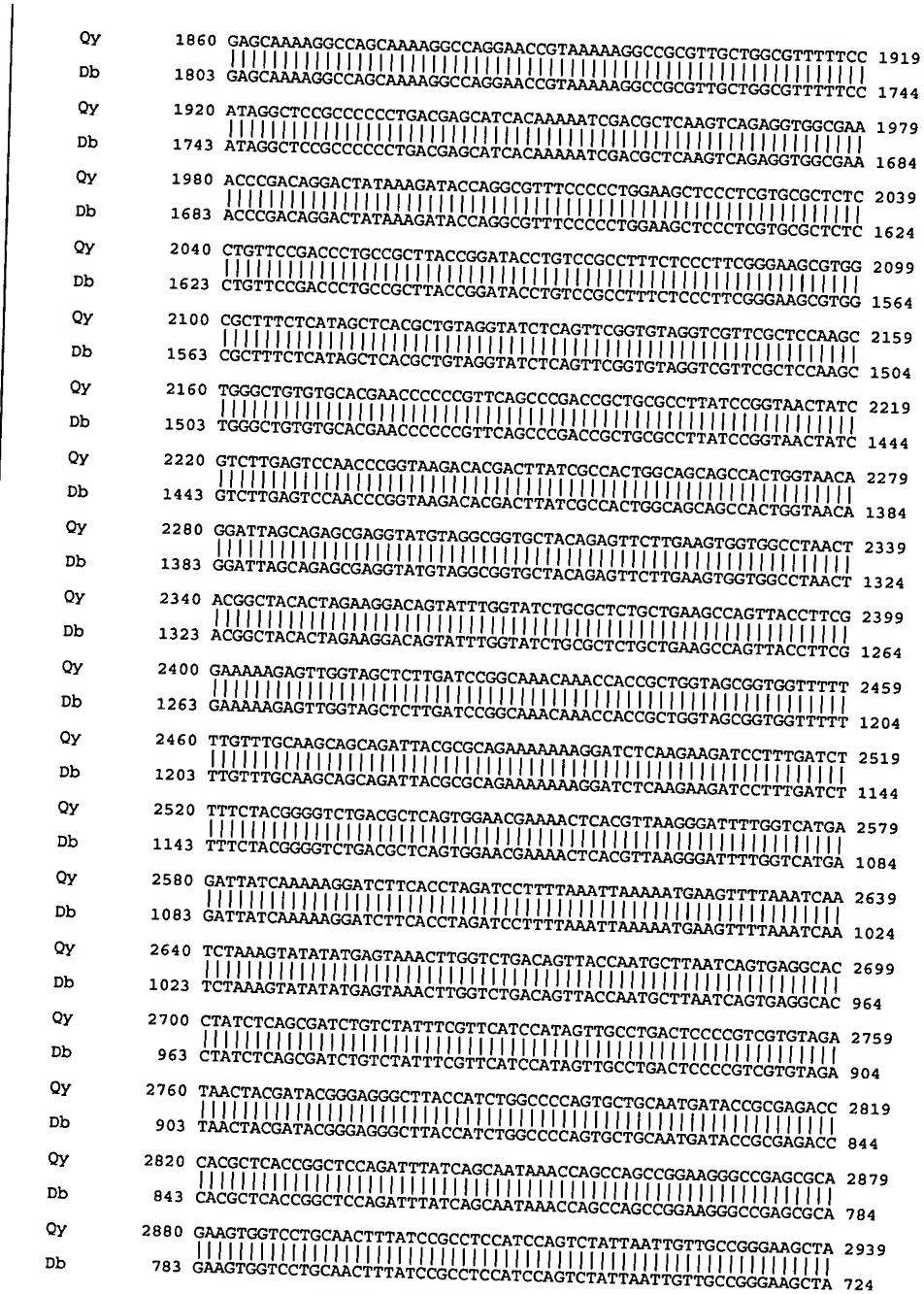
```
Db        903 TAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACC 844

Qy       2820 CACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCA 2879

Db        843 CACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCA 784

Qy       2880 GAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTA 2939

Db        783 GAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTA 724

Qy       2940 GAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCG 2999

Db        723 GAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCG 664

Qy       3000 TGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGC 3059

Db        663 TGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGC 604

Qy       3060 GAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCG 3119

Db        603 GAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCG 544

Qy       3120 TTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATT 3179

Db        543 TTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATT 484

Qy       3180 CTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGT 3239

Db        483 CTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGT 424

Qy       3240 CATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATA 3299

Db        423 CATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATA 364

Qy       3300 ATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGC 3359

Db        363 ATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGC 304

Qy       3360 GAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCAC 3419

Db        303 GAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCAC 244

Qy       3420 CCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAA 3479

Db        243 CCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAA 184

Qy       3480 GGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCT 3539

Db        183 GGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCT 124

Qy       3540 TCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATAT 3599

Db        123 TCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATAT 64

Qy       3600 TTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGC 3659

Db          63 TTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGC 4

Qy       3660 CAC 3662

Db           3 CAC 1
```

RESULT 5
US-10-684-134-32/c
; Sequence 32, Application US/10684134
; Publication No. US20040142433A1
; GENERAL INFORMATION:
; APPLICANT: Large Scale Biology Corporation
; TITLE OF INVENTION: POPULATION OF POLYNUCLEOTIDE SEQUENCE VARIANTS
; FILE REFERENCE: LSBC-PLG5392-CIP
; CURRENT APPLICATION NUMBER: US/10/684,134
; CURRENT FILING DATE: 2003-10-10
; PRIOR APPLICATION NUMBER: 60/402,342
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 10/066,390
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: 60/268,785
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: 60/266,386
; PRIOR FILING DATE: 2002-02-02
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 32
; LENGTH: 3637
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Encodes Cycle 3 BFP gene
US-10-684-134-32

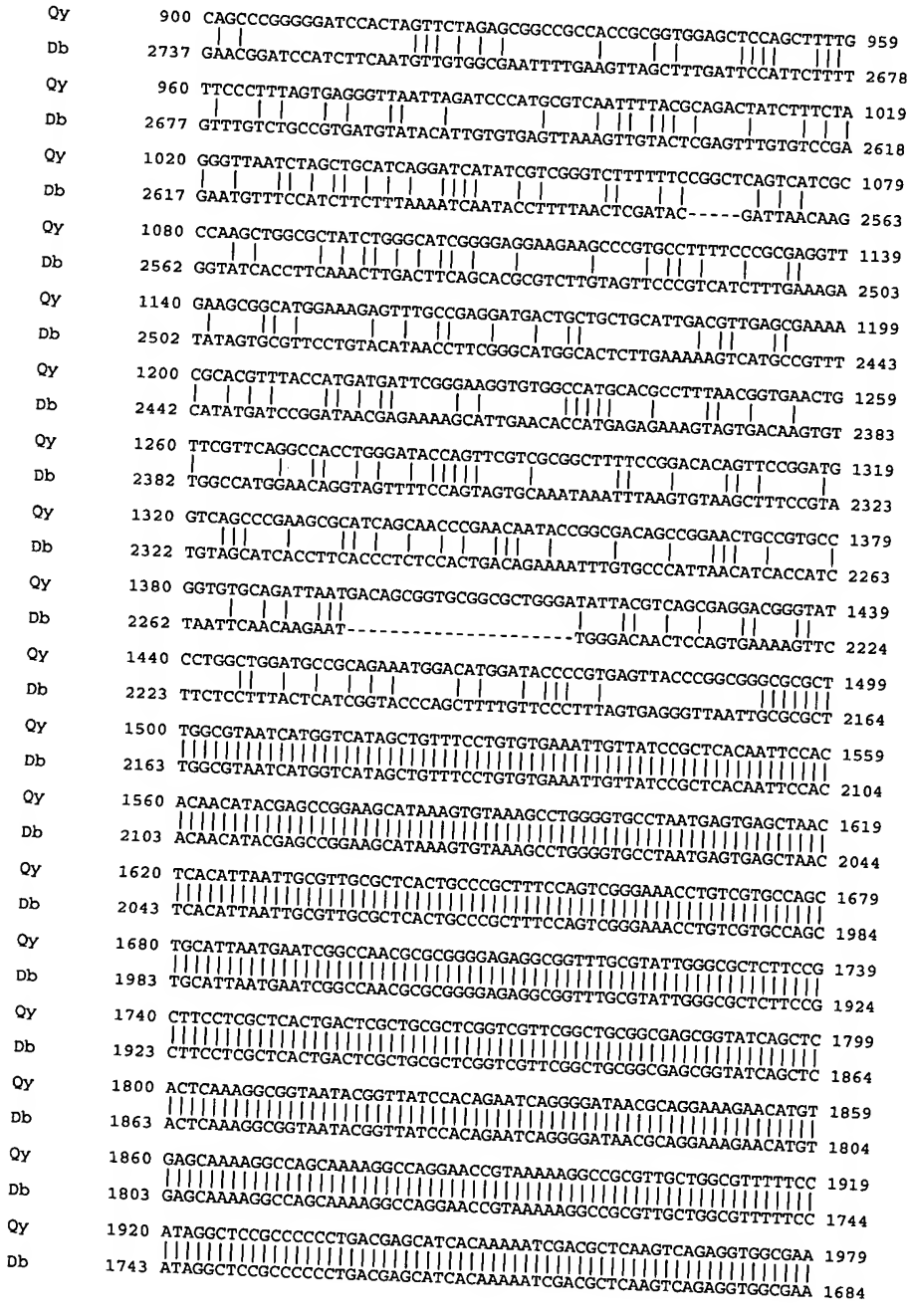
Query Match 72.8%; Score 2665.4; DB 17; Length 3637;
Best Local Similarity 83.7%; Pred. No. 0;
Matches 3065; Conservative 0; Mismatches 571; Indels 27; Gaps 3;

```
Qy          1 CTAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTC 60

Db       3637 CTAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTC 3578

Qy         61 ATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGA 120

Db       3577 ATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGA 3518

Qy        121 GATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTC 180

Db       3517 GATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTC 3458

Qy        181 CAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACC 240

Db       3457 CAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACC 3398

Qy        241 CTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAG 300

Db       3397 CTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAG 3338

Qy        301 CCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAA 360

Db       3337 CCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAA 3278

Qy        361 AGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCAC 420

Db       3277 AGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCAC 3218

Qy        421 CACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCCCATTCGCCATTCAGGCTGCG 480

Db       3217 CACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCCCATTCGCCATTCAGGCTGCG 3158

Qy        481 CAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGG 540

Db       3157 CAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGG 3098

Qy        541 GGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTG 600

Db       3097 GGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTG 3038

Qy        601 TAAAACGACGGCCAGTGAGCGCGCCTCGTTCATTCACGTTTTTGAACCCGTGGAGGACGG 660

Db       3037 TAAAACGACGGCCAGTGAGCGCGCGTAATACGACTCACTATAGGGCGAATTGGAGCTCCA 2978

Qy        661 GCAGACTCGCGGTGCAAATGTGTTTTACAGCGTGATGGAGCAGATGAAGATGCTCGACAC 720

Db       2977 CCGCGGTGGCGGCCGCTCTAGAACTAGTGGATCCCCCGGGCTGCAGGAATTCTTATTTGT 2918

Qy        721 GCTGCAGAACACGCAGCTAGATTAACCCTAGAAAGATAATCATATTGTGACGTACGTTAA 780
```





```
Qy       2940 GAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCG 2999
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        723 GAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCG 664

Qy       3000 TGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGC 3059
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        663 TGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGC 604

Qy       3060 GAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCG 3119
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        603 GAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCG 544

Qy       3120 TTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATT 3179
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        543 TTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATT 484

Qy       3180 CTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGT 3239
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        483 CTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGT 424

Qy       3240 CATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATA 3299
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        423 CATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATA 364

Qy       3300 ATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGC 3359
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 ATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGC 304

Qy       3360 GAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCAC 3419
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 GAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCAC 244

Qy       3420 CCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAA 3479
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 CCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAA 184

Qy       3480 GGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCT 3539
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 GGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCT 124

Qy       3540 TCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATAT 3599
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 TCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATAT 64

Qy       3600 TTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGC 3659
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          63 TTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGC 4

Qy       3660 CAC 3662
             |||
Db          3 CAC 1


RESULT 6
US-10-637-758-32/c
; Sequence 32, Application US/10637758
; Publication No. US20040180352A1
; GENERAL INFORMATION:
;  APPLICANT: Large Scale Biology Corporation
;  TITLE OF INVENTION: POPULATION OF POLYNUCLEOTIDE SEQUENCE VARIANTS
;  FILE REFERENCE: LSBC-PLG5392-CIP
;  CURRENT APPLICATION NUMBER: US/10/637,758
;  CURRENT FILING DATE:  2003-08-08
;  PRIOR APPLICATION NUMBER: US/10/280,913
;  PRIOR FILING DATE: 2002-10-25
;  PRIOR APPLICATION NUMBER: 60/402,342
;  PRIOR FILING DATE: 2002-08-08
;  PRIOR APPLICATION NUMBER: 10/066,390
;  PRIOR FILING DATE: 2002-02-01
;  PRIOR APPLICATION NUMBER: 60/268,785
;  PRIOR FILING DATE: 2001-02-14
;  PRIOR APPLICATION NUMBER: 60/266,386
;  PRIOR FILING DATE: 2002-02-02
;  NUMBER OF SEQ ID NOS: 35
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 32
;   LENGTH: 3637
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Encodes Cycle 3 BFP gene
US-10-637-758-32

  Query Match             72.8%;  Score 2665.4;  DB 17;  Length 3637;
  Best Local Similarity   83.7%;  Pred. No. 0;
  Matches 3065;  Conservative    0;  Mismatches  571;  Indels   27;  Gaps    3;

Qy          1 CTAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3637 CTAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTC 3578

Qy         61 ATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3577 ATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGA 3518

Qy        121 GATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3517 GATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTC 3458

Qy        181 CAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3457 CAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACC 3398

Qy        241 CTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3397 CTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAG 3338

Qy        301 CCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3337 CCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAA 3278

Qy        361 AGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCAC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3277 AGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCAC 3218

Qy        421 CACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCCCATTCGCCATTCAGGCTGCG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3217 CACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCCCATTCGCCATTCAGGCTGCG 3158

Qy        481 CAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3157 CAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGG 3098

Qy        541 GGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3097 GGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTG 3038

Qy        601 TAAAACGACGGCCAGTGAGCGCGCCTCGTTCATTCACGTTTTTGAACCCGTGGAGGACGG 660
              |||||||||||||||||||||||| | | |     | | |     |||||  |
Db       3037 TAAAACGACGGCCAGTGAGCGCGCGTAATACGACTCACTATAGGGCGAATTGGAGCTCCA 2978

Qy        661 GCAGACTCGCGGTGCAAATGTGTTTTACAGCGTGATGGAGCAGATGAAGATGCTCGACAC 720
              |    | ||||      |     |     |     | || | | | | |  |
Db       2977 CCGCGGTGGCGGCCGCTCTAGAACTAGTGGATCCCCCGGGCTGCAGGAATTCTTATTTGT 2918

Qy        721 GCTGCAGAACACGCAGCTAGATTAACCCTAGAAAGATAATCATATTGTGACGTACGTTAA 780
              |   | |       | | | |||   |   | | | |   || | ||  |
Db       2917 ATAGTTCATCCATGCCATGTGTAATCCCAGCAGCAGTTACAAACTCAAGAAGGACCATGT 2858

Qy        781 AGATAATC-ATGCGTAAAATTGACGCATGGGATCTGTAATACGACTCACTATAGGGCGAA 839
              |  |  | | |||     |   |   | | |          ||     | ||
Db       2857 GGTCACGCTTTTCGTTGGGATCTTTCGAAAGGGCAGATTGTGTCGACAGGTAATGGTTGT 2798

Qy        840 TTGGGTACCGGGCCCCCCCTCGAGGTCGACGGTATCGATAAGCTTGATATCGAATTCCTG 899
              |||  |  || |  ||          |       |  ||     |  |
Db       2797 CTGGTAAAAGGACAGGGCCATCGCCAATTGGAGTATTTTGTTGATAATGGTCTGCTAGTT 2738
```

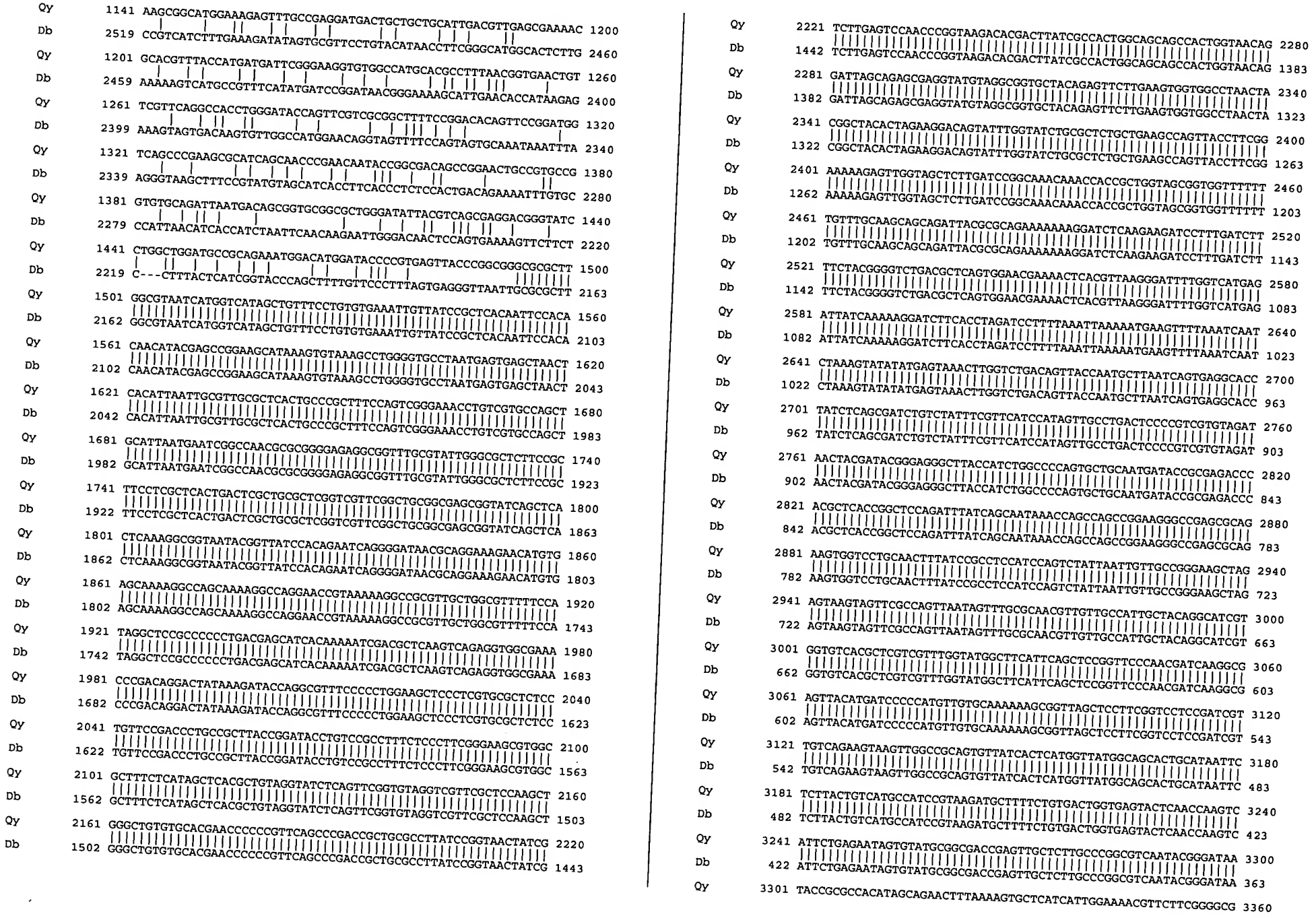

Qy 900 CAGCCCGGGGGATCCACTAGTTCTAGAGCGGCCGCCACCGCGGTGGAGCTCCAGCTTTTG 959
Db 2737 GAACGGATCCATCTTCAATGTTGTGGCGAATTTTGAAGTTAGCTTTGATTCCATTCTTTT 2678
Qy 960 TTCCCTTTAGTGAGGGTTAATTAGATCCCATGCGTCAATTTTACGCAGACTATCTTTCTA 1019
Db 2677 GTTTGTCTGCCGTGATGTATACATTGTGTGAGTTAAAGTTGTACTCGAGTTTGTGTCCGA 2618
Qy 1020 GGGTTAATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGC 1079
Db 2617 GAATGTTTCCATCTTCTTTAAAATCAATACCTTTTAACTCGATAC-----GATTAACAAG 2563
Qy 1080 CCAAGCTGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTT 1139
Db 2562 GGTATCACCTTCAAACTTGACTTCAGCACGCGTCTTGTAGTTCCCGTCATCTTTGAAAGA 2503
Qy 1140 GAAGCGGCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAA 1199
Db 2502 TATAGTGCGTTCCTGTACATAACCTTCGGGCATGGCACTCTTGAAAAAGTCATGCCGTTT 2443
Qy 1200 CGCACGTTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTG 1259
Db 2442 CATATGATCCGGATAACGAGAAAAGCATTGAACACCATGAGAGAAAGTAGTGACAAGTGT 2383
Qy 1260 TTCGTTCAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATG 1319
Db 2382 TGGCCATGGAACAGGTAGTTTTCCAGTAGTGCAAATAAATTTAAGTGTAAGCTTTCCGTA 2323
Qy 1320 GTCAGCCCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCC 1379
Db 2322 TGTAGCATCACCTTCACCCTCTCCACTGACAGAAAATTTGTGCCCATTAACATCACCATC 2263
Qy 1380 GGTGTGCAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTAT 1439
Db 2262 TAATTCAACAAGAAT----------TGGGACAACTCCAGTGAAAAGTTC 2224
Qy 1440 CCTGGCTGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCT 1499
Db 2223 TTCTCCTTTACTCATCGGTACCCAGCTTTTGTTCCCTTTAGTGAGGGTTAATTGCGCGCT 2164
Qy 1500 TGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCAC 1559
Db 2163 TGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCAC 2104
Qy 1560 ACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAAC 1619
Db 2103 ACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAAC 2044
Qy 1620 TCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGC 1679
Db 2043 TCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGC 1984
Qy 1680 TGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCG 1739
Db 1983 TGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCG 1924
Qy 1740 CTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTC 1799
Db 1923 CTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTC 1864
Qy 1800 ACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGT 1859
Db 1863 ACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGT 1804
Qy 1860 GAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCC 1919
Db 1803 GAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCC 1744
Qy 1920 ATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAA 1979
Db 1743 ATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAA 1684
Qy 1980 ACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTC 2039
Db 1683 ACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTC 1624
Qy 2040 CTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGG 2099
Db 1623 CTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGG 1564
Qy 2100 CGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGC 2159
Db 1563 CGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGC 1504
Qy 2160 TGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATC 2219
Db 1503 TGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATC 1444
Qy 2220 GTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACA 2279
Db 1443 GTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACA 1384
Qy 2280 GGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACT 2339
Db 1383 GGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACT 1324
Qy 2340 ACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCG 2399
Db 1323 ACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCG 1264
Qy 2400 GAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTT 2459
Db 1263 GAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTT 1204
Qy 2460 TTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCT 2519
Db 1203 TTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCT 1144
Qy 2520 TTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGA 2579
Db 1143 TTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGA 1084
Qy 2580 GATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAA 2639
Db 1083 GATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAA 1024
Qy 2640 TCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCAC 2699
Db 1023 TCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCAC 964
Qy 2700 CTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGA 2759
Db 963 CTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGA 904
Qy 2760 TAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACC 2819
Db 903 TAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACC 844
Qy 2820 CACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCA 2879
Db 843 CACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCA 784
Qy 2880 GAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTA 2939
Db 783 GAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTA 724
Qy 2940 GAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCG 2999
Db 723 GAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCG 664
Qy 3000 TGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGC 3059
Db 663 TGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGC 604
Qy 3060 GAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCG 3119

```
Db        603 GAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCG 544

Qy       3120 TTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATT 3179

Db        543 TTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATT 484

Qy       3180 CTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGT 3239

Db        483 CTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGT 424

Qy       3240 CATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATA 3299

Db        423 CATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATA 364

Qy       3300 ATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGC 3359

Db        363 ATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGC 304

Qy       3360 GAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCAC 3419

Db        303 GAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCAC 244

Qy       3420 CCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAA 3479

Db        243 CCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAA 184

Qy       3480 GGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCT 3539

Db        183 GGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCT 124

Qy       3540 TCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATAT 3599

Db        123 TCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATAT 64

Qy       3600 TTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGC 3659

Db          63 TTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGC 4

Qy       3660 CAC 3662

Db          3 CAC 1
```

RESULT 7
US-10-066-390-4/c
; Sequence 4, Application US/10066390
; Publication No. US20020146732A1
; GENERAL INFORMATION:
; APPLICANT: Padgett, Hal S.
; APPLICANT: Lindbo, John A.
; APPLICANT: Fitzmaurice, Wayne P.
; TITLE OF INVENTION: A Method of Increasing Complementarity
; TITLE OF INVENTION: In A Heteroduplex
; FILE REFERENCE: P-LG 4878
; CURRENT APPLICATION NUMBER: US/10/066,390
; CURRENT FILING DATE: 2002-02-01
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 3637
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic construct
US-10-066-390-4

Query Match 72.8%; Score 2664.8; DB 13; Length 3637;
Best Local Similarity 83.8%; Pred. No. 0;
Matches 3070; Conservative 0; Mismatches 567; Indels 25; Gaps 4;

```
Qy          1 CTAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTC 60

Db       3637 CTAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTC 3578

Qy         61 ATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGA 120

Db       3577 ATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGA 3518

Qy        121 GATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTC 180

Db       3517 GATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTC 3458

Qy        181 CAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACC 240

Db       3457 CAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACC 3398

Qy        241 CTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAG 300

Db       3397 CTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAG 3338

Qy        301 CCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAA 360

Db       3337 CCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAA 3278

Qy        361 AGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCAC 420

Db       3277 AGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCAC 3218

Qy        421 CACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCCCATTCGCCATTCAGGCTGCG 480

Db       3217 CACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCCCATTCGCCATTCAGGCTGCG 3158

Qy        481 CAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGG 540

Db       3157 CAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGG 3098

Qy        541 GGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTG 600

Db       3097 GGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTG 3038

Qy        601 TAAAACGACGGCCAGTGAGCGCGCCTCGTTCATTCACGTTTTTGAACCCGTGGAGGACGG 660

Db       3037 TAAAACGACGGCCAGTGAGCGCGCGTAATACGACTCACTATAGGGCGAATTGGAGCTCCA 2978

Qy        661 GCAGACTCGCGGTGCAAATGTGTTTTACAGCGTGATGGAGCAGATGAAGATGCTCGACAC 720

Db       2977 CCGCGGTGGCGG----------CCGCTCTAGAACTAGTGGATCCCC--CGG 2939

Qy        721 GCTGCAGAACACGCAGCTAGATTAACCCTAGAAAGATAATCATATTGTGACGTACGTTAA 780

Db       2938 GCTGCAGGAATTCTTATTTGTATAGTTCATCCATGCCATGTGTAATCCCAGCAGCAGTTA 2879

Qy        781 AGATAATCATGCGTAAAATTGACGCATGGGATCTGTAATACGACTCACTATAGGGCGAAT 840

Db       2878 CAAACTCAAGAAGGACCATGTGGTCACGCTTTTCGTTGGGATCTTTCGAAAGGGCAGATT 2819

Qy        841 TGGGTACCGGGCCCCCCCTCGAGGTCGACGGTATCGATAAGCTTGATATCGAATTCCTGC 900

Db       2818 GTGTCGACAGGTAATGGTTGTCTGGTAAAAGGACAGGGCCATCGCCAATTGGAGTATTTT 2759

Qy        901 AGCCCGGGGGATCCACTAGTTCTAGAGCGGCCGCCACCGCGGTGGAGCTCCAGCTTTTGT 960

Db       2758 GTTGATAATGGTCTGCTAGTTGAACGGATCCATCTTCAATGTTGTGGCGAATTTTGAAGT 2699

Qy        961 TCCCTTTAGTGAGGGTTAATTAGATCCCATGCGTCAATTTTACGCAGACTATCTTTCTAG 1020

Db       2698 TAGCTTTGATTCCATTCTTTTGTTTGTCTGCCGTGATGTATACATTGTGTGAG-TTATAG 2640

Qy       1021 GGTTAATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCC 1080

Db       2639 TTGTACTCGAGTTTGTGTCCGAGAATGTTTCCATCTTCTTTAAAATCAATACCTTTTAAC 2580

Qy       1081 CAAGCTGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTG 1140

Db       2579 TCGATACGATTAACAAGGGTATCACCTTCAAACTTGACTTCAGCACGCGTCTTGTAGTTC 2520
```

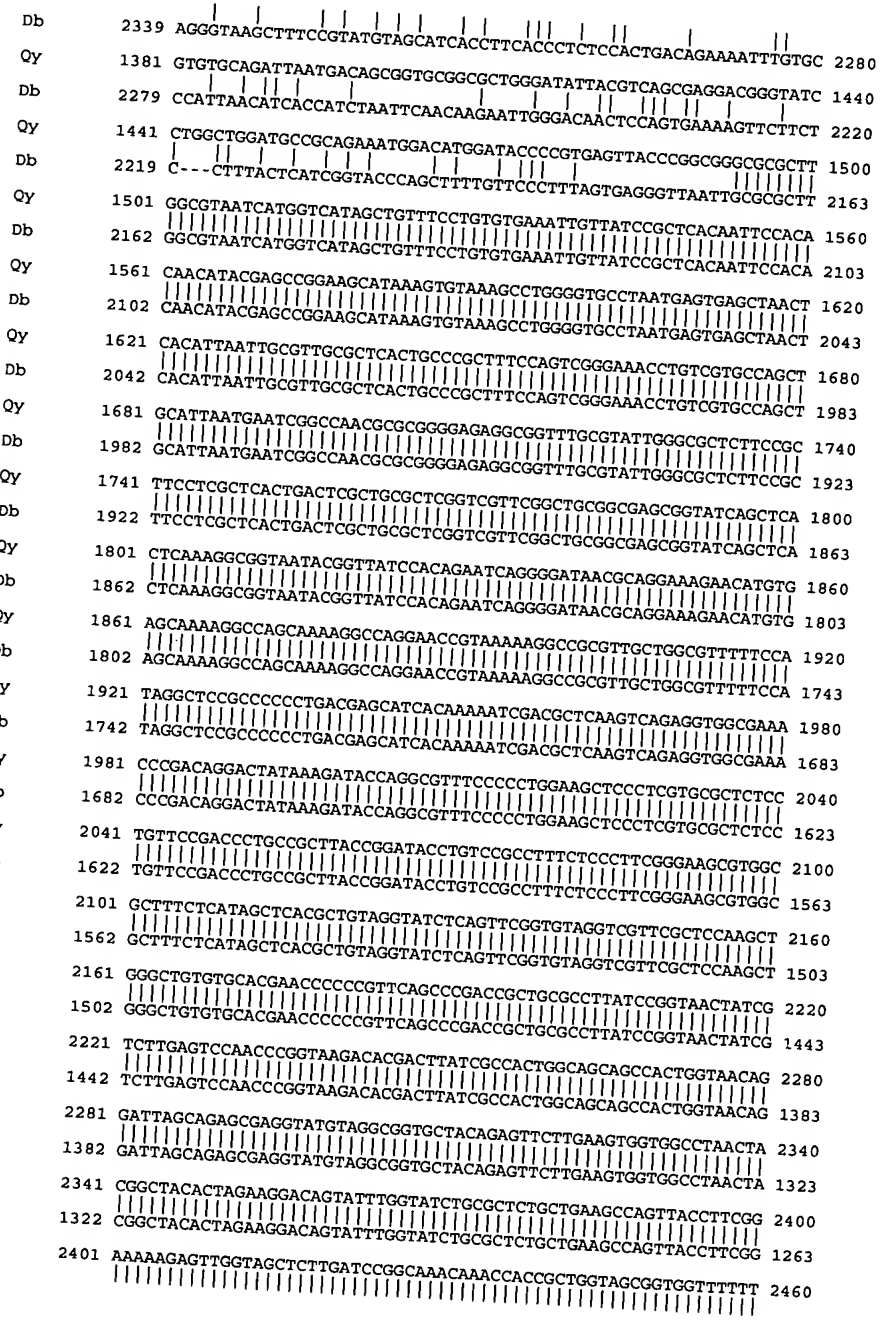

Qy 1141 AAGCGGCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAAC 1200
Db 2519 CCGTCATCTTTGAAAGATATAGTGCGTTCCTGTACATAACCTTCGGGCATGGCACTCTTG 2460
Qy 1201 GCACGTTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGT 1260
Db 2459 AAAAAGTCATGCCGTTTCATATGATCCGGATAACGGGAAAAGCATTGAACACCATAAGAG 2400
Qy 1261 TCGTTCAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGG 1320
Db 2399 AAAGTAGTGACAAGTGTTGGCCATGGAACAGGTAGTTTTCCAGTAGTGCAAATAAATTTA 2340
Qy 1321 TCAGCCCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCG 1380
Db 2339 AGGGTAAGCTTTCCGTATGTAGCATCACCTTCACCCTCTCCACTGACAGAAAATTTGTGC 2280
Qy 1381 GTGTGCAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATC 1440
Db 2279 CCATTAACATCACCATCTAATTCAACAAGAATTGGGACAACTCCAGTGAAAAGTTCTTCT 2220
Qy 1441 CTGGCTGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCTT 1500
Db 2219 C---CTTTACTCATCGGTACCCAGCTTTGTTCCCTTTAGTGAGGGTTAATTGCGCGCTT 2163
Qy 1501 GGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACA 1560
Db 2162 GGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACA 2103
Qy 1561 CAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACT 1620
Db 2102 CAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACT 2043
Qy 1621 CACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCT 1680
Db 2042 CACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCT 1983
Qy 1681 GCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGC 1740
Db 1982 GCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGC 1923
Qy 1741 TTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCA 1800
Db 1922 TTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCA 1863
Qy 1801 CTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTG 1860
Db 1862 CTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTG 1803
Qy 1861 AGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCA 1920
Db 1802 AGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCA 1743
Qy 1921 TAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAA 1980
Db 1742 TAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAA 1683
Qy 1981 CCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCC 2040
Db 1682 CCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCC 1623
Qy 2041 TGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGC 2100
Db 1622 TGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGC 1563
Qy 2101 GCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCT 2160
Db 1562 GCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCT 1503
Qy 2161 GGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCG 2220
Db 1502 GGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCG 1443
Qy 2221 TCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAG 2280
Db 1442 TCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAG 1383
Qy 2281 GATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTA 2340
Db 1382 GATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTA 1323
Qy 2341 CGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGG 2400
Db 1322 CGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGG 1263
Qy 2401 AAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTT 2460
Db 1262 AAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTT 1203
Qy 2461 TGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTT 2520
Db 1202 TGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTT 1143
Qy 2521 TTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAG 2580
Db 1142 TTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAG 1083
Qy 2581 ATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAAT 2640
Db 1082 ATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAAT 1023
Qy 2641 CTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACC 2700
Db 1022 CTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACC 963
Qy 2701 TATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGAT 2760
Db 962 TATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGAT 903
Qy 2761 AACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCC 2820
Db 902 AACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCC 843
Qy 2821 ACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAG 2880
Db 842 ACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAG 783
Qy 2881 AAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAG 2940
Db 782 AAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAG 723
Qy 2941 AGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGT 3000
Db 722 AGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGT 663
Qy 3001 GGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCG 3060
Db 662 GGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCG 603
Qy 3061 AGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGT 3120
Db 602 AGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGT 543
Qy 3121 TGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTC 3180
Db 542 TGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTC 483
Qy 3181 TCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTC 3240
Db 482 TCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTC 423
Qy 3241 ATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAA 3300
Db 422 ATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAA 363
Qy 3301 TACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCG 3360

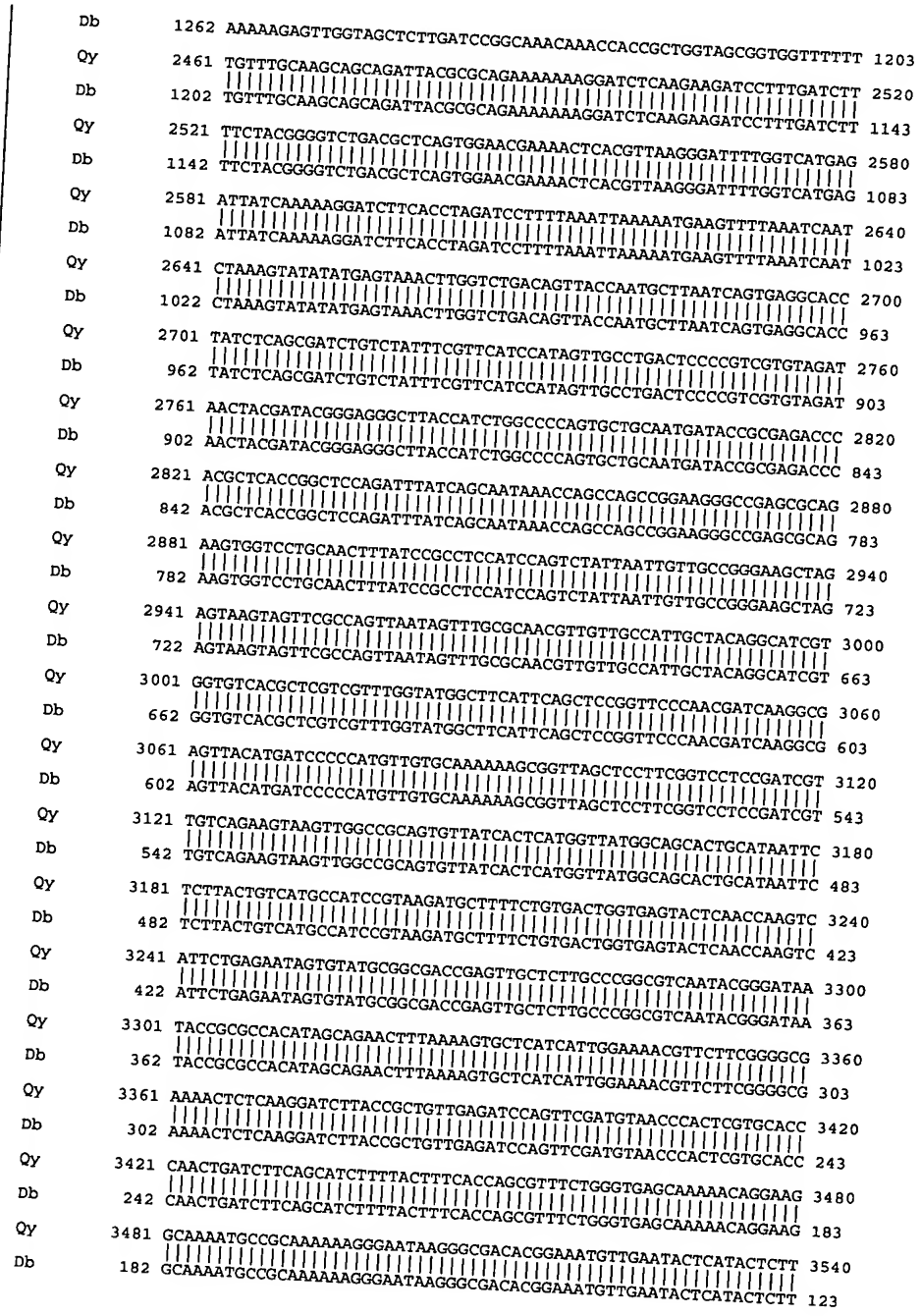
```
Db        362 TACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCG 303

Qy       3361 AAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACC 3420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 AAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACC 243

Qy       3421 CAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAG 3480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 CAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAG 183

Qy       3481 GCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTT 3540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 GCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTT 123

Qy       3541 CCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATT 3600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 CCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATT 63

Qy       3601 TGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCC 3660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          62 TGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCC 3

Qy       3661 AC 3662
              ||
Db          2 AC 1
```

RESULT 8
US-10-206-030-4/c
; Sequence 4, Application US/10206030
; Publication No. US20020177160A1
; GENERAL INFORMATION:
; APPLICANT: Padgett, Hal S.
; APPLICANT: Lindbo, John A.
; APPLICANT: Fitzmaurice, Wayne P.
; TITLE OF INVENTION: A Method of Increasing Complementarity
; TITLE OF INVENTION: In A Heteroduplex
; FILE REFERENCE: P-LG 4878
; CURRENT APPLICATION NUMBER: US/10/206,030
; CURRENT FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: US/10/066,390
; PRIOR FILING DATE: 2002-02-01
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 3637
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic construct
US-10-206-030-4

Query Match 72.8%; Score 2664.8; DB 13; Length 3637;
Best Local Similarity 83.8%; Pred. No. 0;
Matches 3070; Conservative 0; Mismatches 567; Indels 25; Gaps 4;

```
Qy          1 CTAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3637 CTAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTC 3578

Qy         61 ATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3577 ATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGA 3518

Qy        121 GATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3517 GATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTC 3458

Qy        181 CAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3457 CAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACC 3398

Qy        241 CTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3397 CTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAG 3338

Qy        301 CCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3337 CCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAA 3278

Qy        361 AGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCAC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3277 AGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCAC 3218

Qy        421 CACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCCCATTCGCCATTCAGGCTGCG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3217 CACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCCCATTCGCCATTCAGGCTGCG 3158

Qy        481 CAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3157 CAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGG 3098

Qy        541 GGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3097 GGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTG 3038

Qy        601 TAAAACGACGGCCAGTGAGCGCGCCTCGTTCATTCACGTTTTTGAACCCGTGGAGGACGG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3037 TAAAACGACGGCCAGTGAGCGCGCGTAATACGACTCACTATAGGGCGAATTGGAGCTCCA 2978

Qy        661 GCAGACTCGCGGTGCAAATGTGTTTTACAGCGTGATGGAGCAGATGAAGATGCTCGACAC 720
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2977 CCGCGGTGGCGG----------CGCTCTAGAACTAGTGGATCCCC--CGG 2939

Qy        721 GCTGCAGAACACGCAGCTAGATTAACCCTAGAAAGATAATCATATTGTGACGTACGTTAA 780
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2938 GCTGCAGGAATTCTTATTTGTATAGTTCATCCATGCCATGTGTAATCCCAGCAGCAGTTA 2879

Qy        781 AGATAATCATGCGTAAAATTGACGCATGGGATCTGTAATACGACTCACTATAGGGCGAAT 840
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2878 CAAACTCAAGAAGGACCATGTGGTCACGCTTTTCGTTGGGATCTTTCGAAAGGGCAGATT 2819

Qy        841 TGGGTACCGGGCCCCCCCTCGAGGTCGACGGTATCGATAAGCTTGATATCGAATTCCTGC 900
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2818 GTGTCGACAGGTAATGGTTGTCTGGTAAAAGGACAGGGCCATCGCCAATTGGAGTATTTT 2759

Qy        901 AGCCCGGGGGATCCACTAGTTCTAGAGCGGCCGCCACCGCGGTGGAGCTCCAGCTTTTGT 960
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2758 GTTGATAATGGTCTGCTAGTTGAACGGATCCATCTTCAATGTTGTGGCGAATTTTGAAGT 2699

Qy        961 TCCCTTTAGTGAGGGTTAATTAGATCCCATGCGTCAATTTTACGCAGACTATCTTTCTAG 1020
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2698 TAGCTTTGATTCCATTCTTTTGTTTGTCTGCCGTGATGTATACATTGTGTGAG-TTATAG 2640

Qy       1021 GGTTAATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCC 1080
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2639 TTGTACTCGAGTTTGTGTCCGAGAATGTTTCCATCTTCTTTAAAATCAATACCTTTTAAC 2580

Qy       1081 CAAGCTGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTG 1140
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2579 TCGATACGATTAACAAGGGTATCACCTTCAAACTTGACTTCAGCACGCGTCTTGTAGTTC 2520

Qy       1141 AAGCGGCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAAC 1200
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2519 CCGTCATCTTTGAAAGATATAGTGCGTTCCTGTACATAACCTTCGGGCATGGCACTCTTG 2460

Qy       1201 GCACGTTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGT 1260
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2459 AAAAAGTCATGCCGTTTCATATGATCCGGATAACGGGAAAAGCATTGAACACCATAAGAG 2400

Qy       1261 TCGTTCAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGG 1320
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2399 AAAGTAGTGACAAGTGTTGGCCATGGAACAGGTAGTTTTCCAGTAGTGCAAATAAATTTA 2340

Qy       1321 TCAGCCCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCG 1380
```





```
Qy       3541 CCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATT 3600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 CCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATT 63

Qy       3601 TGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCC 3660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 TGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCC 3

Qy       3661 AC 3662
              ||
Db          2 AC 1


RESULT 9
US-10-211-079-4/c
; Sequence 4, Application US/10211079
; Publication No. US20030148315A1
; GENERAL INFORMATION:
;  APPLICANT: Padgett, Hal S.
;  APPLICANT:  Vaewhongs, Andrew A.
;  APPLICANT:  Vojdani, Fakhrieh S.
;  APPLICANT:  Smith, Mark L.
;  TITLE OF INVENTION: Nucleic Acid Molecules Encoding CEL I
;  TITLE OF INVENTION:  Endonuclease and Methods of Use Thereof
;  FILE REFERENCE: P-LG 5381
;  CURRENT APPLICATION NUMBER: US/10/211,079
;  CURRENT FILING DATE:  2002-08-01
;  PRIOR APPLICATION NUMBER: US 10/098,155
;  PRIOR FILING DATE: 2002-03-14
;  NUMBER OF SEQ ID NOS: 29
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
;   LENGTH: 3637
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: synthetic construct
US-10-211-079-4

  Query Match             72.8%;  Score 2664.8;  DB 15;  Length 3637;
  Best Local Similarity   83.8%;  Pred. No. 0;
  Matches 3070;  Conservative    0;  Mismatches  567;  Indels   25;  Gaps    4;

Qy          1 CTAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3637 CTAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTC 3578

Qy         61 ATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3577 ATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGA 3518

Qy        121 GATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3517 GATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTC 3458

Qy        181 CAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3457 CAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACC 3398

Qy        241 CTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3397 CTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAG 3338

Qy        301 CCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3337 CCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAA 3278

Qy        361 AGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCAC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3277 AGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCAC 3218

Qy        421 CACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCCCATTCGCCATTCAGGCTGCG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3217 CACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCCCATTCGCCATTCAGGCTGCG 3158

Qy        481 CAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3157 CAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGG 3098

Qy        541 GGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3097 GGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTG 3038

Qy        601 TAAAACGACGGCCAGTGAGCGCGCCTCGTTCATTCACGTTTTTGAACCCGTGGAGGACGG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3037 TAAAACGACGGCCAGTGAGCGCGCGTAATACGACTCACTATAGGGCGAATTGGAGCTCCA 2978

Qy        661 GCAGACTCGCGGTGCAAATGTGTTTTACAGCGTGATGGAGCAGATGAAGATGCTCGACAC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2977 CCGCGGTGGCGG----------------------CCGCTCTAGAACTAGTGGATCCCC--CGG 2939

Qy        721 GCTGCAGAACACGCAGCTAGATTAACCCTAGAAAGATAATCATATTGTGACGTACGTTAA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2938 GCTGCAGGAATTCTTATTTGTATAGTTCATCCATGCCATGTGTAATCCCAGCAGCAGTTA 2879

Qy        781 AGATAATCATGCGTAAAATTGACGCATGGGATCTGTAATACGACTCACTATAGGGCGAAT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2878 CAAACTCAAGAAGGACCATGTGGTCACGCTTTTCGTTGGGATCTTTCGAAAGGGCAGATT 2819

Qy        841 TGGGTACCGGGCCCCCCCTCGAGGTCGACGGTATCGATAAGCTTGATATCGAATTCCTGC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2818 GTGTCGACAGGTAATGGTTGTCTGGTAAAAGGACAGGGCCATCGCCAATTGGAGTATTTT 2759

Qy        901 AGCCCGGGGGATCCACTAGTTCTAGAGCGGCCGCCACCGCGGTGGAGCTCCAGCTTTTGT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2758 GTTGATAATGGTCTGCTAGTTGAACGGATCCATCTTCAATGTTGTGGCGAATTTTGAAGT 2699

Qy        961 TCCCTTTAGTGAGGGTTAATTAGATCCCATGCGTCAATTTTACGCAGACTATCTTTCTAG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2698 TAGCTTTGATTCCATTCTTTTGTTTGTCTGCCGTGATGTATACATTGTGTGAG-TTATAG 2640

Qy       1021 GGTTAATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2639 TTGTACTCGAGTTTGTGTCCGAGAATGTTTCCATCTTCTTTAAAATCAATACCTTTTAAC 2580

Qy       1081 CAAGCTGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2579 TCGATACGATTAACAAGGGTATCACCTTCAAACTTGACTTCAGCACGCGTCTTGTAGTTC 2520

Qy       1141 AAGCGGCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAAC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2519 CCGTCATCTTTGAAAGATATAGTGCGTTCCTGTACATAACCTTCGGGCATGGCACTCTTG 2460

Qy       1201 GCACGTTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2459 AAAAAGTCATGCCGTTTCATATGATCCGGATAACGGGAAAAGCATTGAACACCATAAGAG 2400

Qy       1261 TCGTTCAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGG 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2399 AAAGTAGTGACAAGTGTTGGCCATGGAACAGGTAGTTTTCCAGTAGTGCAAATAAATTTA 2340

Qy       1321 TCAGCCCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2339 AGGGTAAGCTTTCCGTATGTAGCATCACCTTCACCCTCTCCACTGACAGAAAATTTGTGC 2280

Qy       1381 GTGTGCAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2279 CCATTAACATCACCATCTAATTCAACAAGAATTGGGACAACTCCAGTGAAAAGTTCTTCT 2220

Qy       1441 CTGGCTGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCTT 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2219 C---CTTTACTCATCGGTACCCAGCTTTTGTTCCCTTTAGTGAGGGTTAATTGCGCGCTT 2163

Qy       1501 GGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       2162 GGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACA 2103

Qy       1561 CAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACT 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2102 CAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACT 2043

Qy       1621 CACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCT 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2042 CACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCT 1983

Qy       1681 GCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGC 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1982 GCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGC 1923

Qy       1741 TTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCA 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1922 TTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCA 1863

Qy       1801 CTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTG 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1862 CTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTG 1803

Qy       1861 AGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCA 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1802 AGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCA 1743

Qy       1921 TAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAA 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1742 TAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAA 1683

Qy       1981 CCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCC 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1682 CCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCC 1623

Qy       2041 TGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGC 2100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1622 TGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGC 1563

Qy       2101 GCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCT 2160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1562 GCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCT 1503

Qy       2161 GGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCG 2220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1502 GGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCG 1443

Qy       2221 TCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAG 2280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1442 TCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAG 1383

Qy       2281 GATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTA 2340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1382 GATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTA 1323

Qy       2341 CGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGG 2400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1322 CGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGG 1263

Qy       2401 AAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTT 2460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1262 AAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTT 1203

Qy       2461 TGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTT 2520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1202 TGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTT 1143

Qy       2521 TTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAG 2580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1142 TTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAG 1083

Qy       2581 ATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAAT 2640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1082 ATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAAT 1023

Qy       2641 CTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACC 2700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1022 CTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACC 963

Qy       2701 TATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGAT 2760
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        962 TATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGAT 903

Qy       2761 AACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCC 2820
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        902 AACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCC 843

Qy       2821 ACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAG 2880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        842 ACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAG 783

Qy       2881 AAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAG 2940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        782 AAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAG 723

Qy       2941 AGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGT 3000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        722 AGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGT 663

Qy       3001 GGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCG 3060
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        662 GGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCG 603

Qy       3061 AGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGT 3120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        602 AGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGT 543

Qy       3121 TGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTC 3180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        542 TGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTC 483

Qy       3181 TCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTC 3240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        482 TCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTC 423

Qy       3241 ATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAA 3300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 ATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAA 363

Qy       3301 TACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCG 3360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 TACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCG 303

Qy       3361 AAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACC 3420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 AAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACC 243

Qy       3421 CAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAG 3480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 CAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAG 183

Qy       3481 GCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTT 3540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 GCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTT 123

Qy       3541 CCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATT 3600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 CCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATT 63

Qy       3601 TGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCC 3660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 TGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCC 3

Qy       3661 AC 3662
              ||
Db          2 AC 1
```

```
RESULT 10
US-10-356-708-17/c
; Sequence 17, Application US/10356708
; Publication No. US20030157682A1
; GENERAL INFORMATION:
;  APPLICANT: Large Scale Biology Corporation
;  TITLE OF INVENTION: MISMATCH ENDONUCLEASES AND METHODS OF USE THEREOF
;  FILE REFERENCE: P-LG 10100
;  CURRENT APPLICATION NUMBER: US/10/356,708
;  CURRENT FILING DATE:  2003-01-31
;  PRIOR APPLICATION NUMBER: 60/353,722
;  PRIOR FILING DATE: 2000-02-01
;  PRIOR APPLICATION NUMBER: 10/098,155
;  PRIOR FILING DATE: 2002-03-14
;  PRIOR APPLICATION NUMBER: 10/211,079
;  PRIOR FILING DATE: 2002-08-01
;  NUMBER OF SEQ ID NOS: 35
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17
;   LENGTH: 3637
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Encodes cycle 3 GFP
US-10-356-708-17

  Query Match             72.8%;  Score 2664.8;  DB 15;  Length 3637;
  Best Local Similarity   83.8%;  Pred. No. 0;
  Matches 3070;  Conservative    0;  Mismatches  567;  Indels   25;  Gaps    4;

Qy          1 CTAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3637 CTAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTC 3578

Qy         61 ATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3577 ATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGA 3518

Qy        121 GATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3517 GATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTC 3458

Qy        181 CAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3457 CAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACC 3398

Qy        241 CTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3397 CTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAG 3338

Qy        301 CCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3337 CCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAA 3278

Qy        361 AGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCAC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3277 AGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCAC 3218

Qy        421 CACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCCCATTCGCCATTCAGGCTGCG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3217 CACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCCCATTCGCCATTCAGGCTGCG 3158

Qy        481 CAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3157 CAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGG 3098

Qy        541 GGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3097 GGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTG 3038

Qy        601 TAAAACGACGGCCAGTGAGCGCGCCTCGTTCATTCACGTTTTTGAACCCGTGGAGGACGG 660
              |||||||||||||||||||||||| | | |  | | |  |||||  |
```

```
Db       3037 TAAAACGACGGCCAGTGAGCGCGCGTAATACGACTCACTATAGGGCGAATTGGAGCTCCA 2978

Qy        661 GCAGACTCGCGGTGCAAATGTGTTTTACAGCGTGATGGAGCAGATGAAGATGCTCGACAC 720
              |  | ||||                |  || |  | ||| | | |
Db       2977 CCGCGGTGGCGG------------------CCGCTCTAGAACTAGTGGATCCCC--CGG 2939

Qy        721 GCTGCAGAACACGCAGCTAGATTAACCCTAGAAAGATAATCATATTGTGACGTACGTTAA 780
              ||||||| |         | |  ||  |    | |  |  || | |  |  | | |
Db       2938 GCTGCAGGAATTCTTATTTGTATAGTTCATCCATGCCATGTGTAATCCCAGCAGCAGTTA 2879

Qy        781 AGATAATCATGCGTAAAATTGACGCATGGGATCTGTAATACGACTCACTATAGGGCGAAT 840
              |   |  | | ||  || |  | ||      |  |  || || | |
Db       2878 CAAACTCAAGAAGGACCATGTGGTCACGCTTTTCGTTGGGATCTTTCGAAAGGGCAGATT 2819

Qy        841 TGGGTACCGGGCCCCCCCTCGAGGTCGACGGTATCGATAAGCTTGATATCGAATTCCTGC 900
              |  | ||      |  |  | | | |          || | | | |
Db       2818 GTGTCGACAGGTAATGGTTGTCTGGTAAAAGGACAGGGCCATCGCCAATTGGAGTATTTT 2759

Qy        901 AGCCCGGGGGATCCACTAGTTCTAGAGCGGCCGCCACCGCGGTGGAGCTCCAGCTTTTGT 960
                  || ||||||  | |  | | | | || ||  |  ||
Db       2758 GTTGATAATGGTCTGCTAGTTGAACGGATCCATCTTCAATGTTGTGGCGAATTTTGAAGT 2699

Qy        961 TCCCTTTAGTGAGGGTTAATTAGATCCCATGCGTCAATTTTACGCAGACTATCTTTCTAG 1020
              | |||| |  |  || | |  ||| | | |  | |  || |||
Db       2698 TAGCTTTGATTCCATTCTTTTGTTTGTCTGCCGTGATGTATACATTGTGTGAG-TTATAG 2640

Qy       1021 GGTTAATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCC 1080
              || || ||  |  || || |  |  | |  |  |||  |  |  |
Db       2639 TTGTACTCGAGTTTGTGTCCGAGAATGTTTCCATCTTCTTTAAAATCAATACCTTTTAAC 2580

Qy       1081 CAAGCTGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTG 1140
                |  |  ||| |||  |  || | |  | | |  ||
Db       2579 TCGATACGATTAACAAGGGTATCACCTTCAAACTTGACTTCAGCACGCGTCTTGTAGTTC 2520

Qy       1141 AAGCGGCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAAC 1200
              |  |  ||  | |  | |  | | |  | |  |  ||
Db       2519 CCGTCATCTTTGAAAGATATAGTGCGTTCCTGTACATAACCTTCGGGCATGGCACTCTTG 2460

Qy       1201 GCACGTTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGT 1260
              |  | |  |  | |  |  |  | | |  |  | || || |||  |
Db       2459 AAAAAGTCATGCCGTTTCATATGATCCGGATAACGGGAAAAGCATTGAACACCATAAGAG 2400

Qy       1261 TCGTTCAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGG 1320
              |  | ||  ||  |  |  |  | |||  |  |  |  |
Db       2399 AAAGTAGTGACAAGTGTTGGCCATGGAACAGGTAGTTTTCCAGTAGTGCAAATAAATTTA 2340

Qy       1321 TCAGCCCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCG 1380
              |  |  |  | | | | | |  |||  |  ||  |  ||
Db       2339 AGGGTAAGCTTTCCGTATGTAGCATCACCTTCACCCTCTCCACTGACAGAAAATTTGTGC 2280

Qy       1381 GTGTGCAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATC 1440
              |  | || |  |  |  |  |  |  ||  ||| || |  |
Db       2279 CCATTAACATCACCATCTAATTCAACAAGAATTGGGACAACTCCAGTGAAAAGTTCTTCT 2220

Qy       1441 CTGGCTGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCTT 1500
              |  ||  |  | | | |  |  | |||  |                ||||||||
Db       2219 C---CTTTACTCATCGGTACCCAGCTTTTGTTCCCTTTAGTGAGGGTTAATTGCGCGCTT 2163

Qy       1501 GGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2162 GGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACA 2103

Qy       1561 CAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACT 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2102 CAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACT 2043

Qy       1621 CACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCT 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2042 CACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCT 1983

Qy       1681 GCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGC 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1982 GCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGC 1923
```

```
Qy       1741 TTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCA 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1922 TTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCA 1863

Qy       1801 CTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTG 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1862 CTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTG 1803

Qy       1861 AGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCA 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1802 AGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCA 1743

Qy       1921 TAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAA 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1742 TAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAA 1683

Qy       1981 CCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCC 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1682 CCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCC 1623

Qy       2041 TGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGC 2100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1622 TGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGC 1563

Qy       2101 GCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCT 2160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1562 GCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCT 1503

Qy       2161 GGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCG 2220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1502 GGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCG 1443

Qy       2221 TCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAG 2280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1442 TCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAG 1383

Qy       2281 GATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTA 2340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1382 GATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTA 1323

Qy       2341 CGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGG 2400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1322 CGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGG 1263

Qy       2401 AAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTT 2460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1262 AAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTT 1203

Qy       2461 TGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTT 2520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1202 TGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTT 1143

Qy       2521 TTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAG 2580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1142 TTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAG 1083

Qy       2581 ATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAAT 2640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1082 ATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAAT 1023

Qy       2641 CTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACC 2700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1022 CTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACC 963

Qy       2701 TATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGAT 2760
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        962 TATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGAT 903

Qy       2761 AACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCC 2820
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        902 AACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCC 843

Qy       2821 ACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAG 2880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        842 ACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAG 783

Qy       2881 AAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAG 2940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        782 AAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAG 723

Qy       2941 AGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGT 3000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        722 AGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGT 663

Qy       3001 GGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCG 3060
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        662 GGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCG 603

Qy       3061 AGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGT 3120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        602 AGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGT 543

Qy       3121 TGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTC 3180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        542 TGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTC 483

Qy       3181 TCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTC 3240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        482 TCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTC 423

Qy       3241 ATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAA 3300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 ATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAA 363

Qy       3301 TACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCG 3360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 TACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCG 303

Qy       3361 AAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACC 3420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 AAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACC 243

Qy       3421 CAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAG 3480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 CAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAG 183

Qy       3481 GCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTT 3540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 GCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTT 123

Qy       3541 CCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATT 3600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 CCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATT 63

Qy       3601 TGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCC 3660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 TGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCC 3

Qy       3661 AC 3662
              ||
Db          2 AC 1
```

RESULT 11
US-10-205-772-4/c
; Sequence 4, Application US/10205772
; Publication No. US20030186261A1
; GENERAL INFORMATION:
; APPLICANT: Padgett, Hal S.
; APPLICANT: Lindbo, John A.
; APPLICANT: Fitzmaurice, Wayne P.
; TITLE OF INVENTION: A Method of Increasing Complementarity
; TITLE OF INVENTION: In A Heteroduplex
; FILE REFERENCE: P-LG 4878
; CURRENT APPLICATION NUMBER: US/10/205,772

```
;  CURRENT FILING DATE:  2002-07-25
;  PRIOR APPLICATION NUMBER: US/10/066,390
;  PRIOR FILING DATE: 2002-02-01
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
;   LENGTH: 3637
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: synthetic construct
US-10-205-772-4

  Query Match             72.8%;  Score 2664.8;  DB 15;  Length 3637;
  Best Local Similarity   83.8%;  Pred. No. 0;
  Matches 3070;  Conservative    0;  Mismatches  567;  Indels   25;  Gaps    4;

Qy          1 CTAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3637 CTAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTC 3578

Qy         61 ATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3577 ATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGA 3518

Qy        121 GATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3517 GATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTC 3458

Qy        181 CAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3457 CAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACC 3398

Qy        241 CTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3397 CTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAG 3338

Qy        301 CCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3337 CCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAA 3278

Qy        361 AGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCAC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3277 AGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCAC 3218

Qy        421 CACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCCCATTCGCCATTCAGGCTGCG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3217 CACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCCCATTCGCCATTCAGGCTGCG 3158

Qy        481 CAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3157 CAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGG 3098

Qy        541 GGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3097 GGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTG 3038

Qy        601 TAAAACGACGGCCAGTGAGCGCGCCTCGTTCATTCACGTTTTTGAACCCGTGGAGGACGG 660
              ||||||||||||||||||||||||||| | | |        | | |        |||| |
Db       3037 TAAAACGACGGCCAGTGAGCGCGCGTAATACGACTCACTATAGGGCGAATTGGAGCTCCA 2978

Qy        661 GCAGACTCGCGGTGCAAATGTGTTTTACAGCGTGATGGAGCAGATGAAGATGCTCGACAC 720
              |   | ||||                  | | || |   |||  | |  |
Db       2977 CCGCGGTGGCGG------------------CCGCTCTAGAACTAGTGGATCCCC--CGG 2939

Qy        721 GCTGCAGAACACGCAGCTAGATTAACCCTAGAAAGATAATCATATTGTGACGTACGTTAA 780
              |||||| |        | |  ||    |  | |    || |    |   | ||
Db       2938 GCTGCAGGAATTCTTATTTGTATAGTTCATCCATGCCATGTGTAATCCCAGCAGCAGTTA 2879

Qy        781 AGATAATCATGCGTAAAATTGACGCATGGGATCTGTAATACGACTCACTATAGGGCGAAT 840
              |  |  | | ||    || |  | ||       |  | ||  || |
Db       2878 CAAACTCAAGAAGGACCATGTGGTCACGCTTTTCGTTGGGATCTTTCGAAAGGGCAGATT 2819

Qy        841 TGGGTACCGGGCCCCCCCTCGAGGTCGACGGTATCGATAAGCTTGATATCGAATTCCTGC 900
              |  | | ||        |  | | | | | |           || | | | |
Db       2818 GTGTCGACAGGTAATGGTTGTCTGGTAAAAGGACAGGGCCATCGCCAATTGGAGTATTTT 2759

Qy        901 AGCCCGGGGGATCCACTAGTTCTAGAGCGGCCGCCACCGCGGTGGAGCTCCAGCTTTTGT 960
                  | ||  |||||| | |  | | | | | ||  ||       |  ||
Db       2758 GTTGATAATGGTCTGCTAGTTGAACGGATCCATCTTCAATGTTGTGGCGAATTTTGAAGT 2699

Qy        961 TCCCTTTAGTGAGGGTTAATTAGATCCCATGCGTCAATTTTACGCAGACTATCTTTCTAG 1020
              | ||||  |    |  ||  |  ||| |  | |||  |  |   || |||
Db       2698 TAGCTTTGATTCCATTCTTTTGTTTGTCTGCCGTGATGTATACATTGTGTGAG-TTATAG 2640

Qy       1021 GGTTAATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCC 1080
              || || ||  |        ||  || |  |  |  |||          |  |  |
Db       2639 TTGTACTCGAGTTTGTGTCCGAGAATGTTTCCATCTTCTTTAAAATCAATACCTTTTAAC 2580

Qy       1081 CAAGCTGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTG 1140
                  |   | ||| |||        ||  | |  |  |  |  |  | |||
Db       2579 TCGATACGATTAACAAGGGTATCACCTTCAAACTTGACTTCAGCACGCGTCTTGTAGTTC 2520

Qy       1141 AAGCGGCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAAC 1200
              |  |  ||  | |  | |      | |  | | |  ||
Db       2519 CCGTCATCTTTGAAAGATATAGTGCGTTCCTGTACATAACCTTCGGGCATGGCACTCTTG 2460

Qy       1201 GCACGTTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGT 1260
              |  | |  |  |  |  |  |  |  |  | || || |||  |
Db       2459 AAAAAGTCATGCCGTTTCATATGATCCGGATAACGGGAAAAGCATTGAACACCATAAGAG 2400

Qy       1261 TCGTTCAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGG 1320
                |  | ||  ||    |  |  | ||| | | |  |  |
Db       2399 AAAGTAGTGACAAGTGTTGGCCATGGAACAGGTAGTTTTCCAGTAGTGCAAATAAATTTA 2340

Qy       1321 TCAGCCCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCG 1380
              |  |  | |  | | |  | |  ||| |  ||  |  ||
Db       2339 AGGGTAAGCTTTCCGTATGTAGCATCACCTTCACCCTCTCCACTGACAGAAAATTTGTGC 2280

Qy       1381 GTGTGCAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATC 1440
              |  | ||  |  |  |  |  | | |  ||  ||| ||  |  |
Db       2279 CCATTAACATCACCATCTAATTCAACAAGAATTGGGACAACTCCAGTGAAAAGTTCTTCT 2220

Qy       1441 CTGGCTGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCTT 1500
              |  ||  | | | | | |  |  | | |  |  |  |||||||
Db       2219 C---CTTTACTCATCGGTACCCAGCTTTGTTCCCTTTAGTGAGGGTTAATTGCGCGCTT 2163

Qy       1501 GGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2162 GGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACA 2103

Qy       1561 CAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACT 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2102 CAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACT 2043

Qy       1621 CACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCT 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2042 CACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCT 1983

Qy       1681 GCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGC 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1982 GCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGC 1923

Qy       1741 TTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCA 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1922 TTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCA 1863

Qy       1801 CTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTG 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1862 CTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTG 1803

Qy       1861 AGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCA 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1802 AGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCA 1743

Qy       1921 TAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAA 1980
```

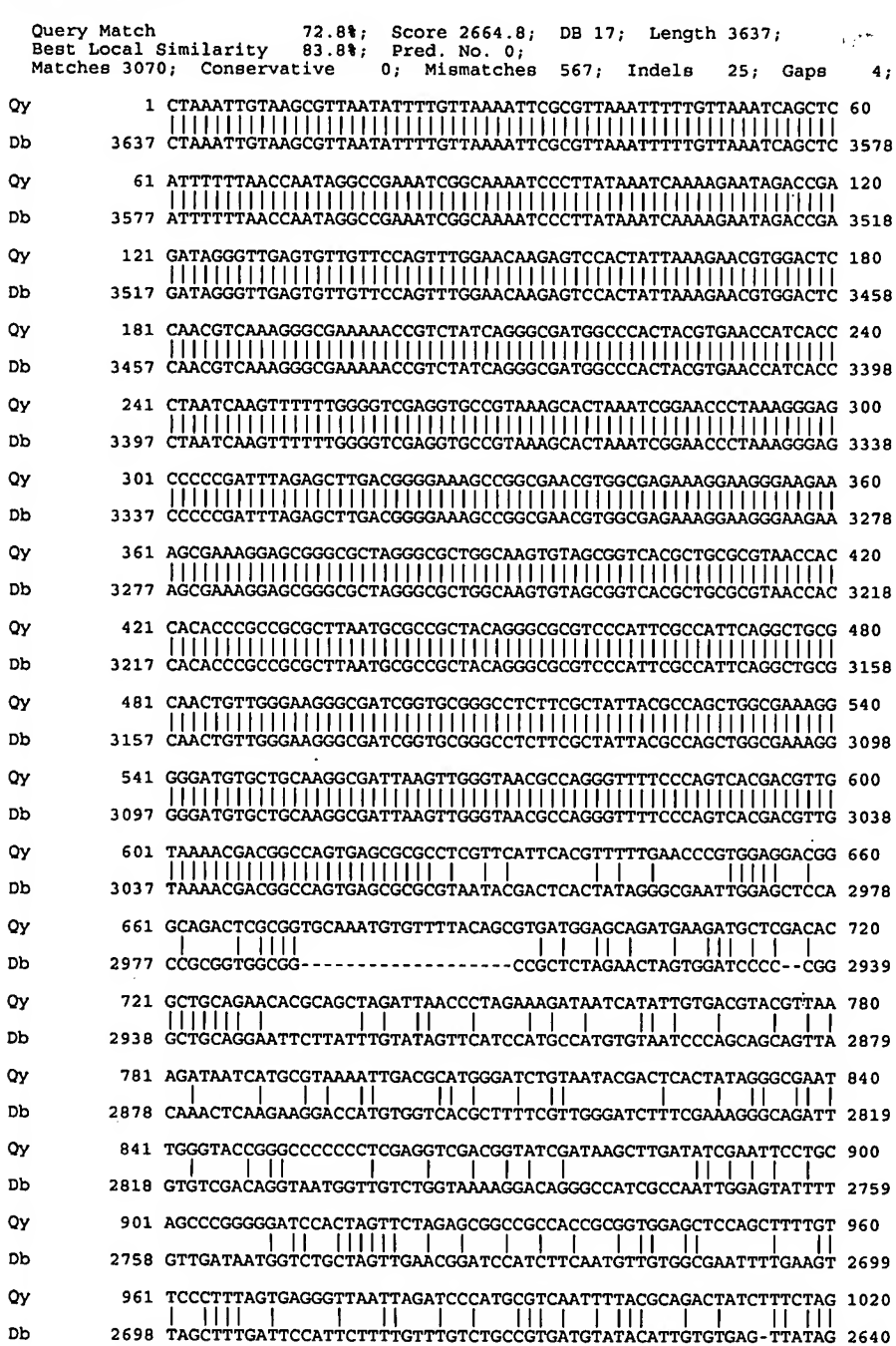
```
Db      1742 TAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAA 1683
Qy      1981 CCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCC 2040
Db      1682 CCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCC 2040
Qy      2041 TGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGC 2100
Db      1622 TGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGC 2100
Qy      2101 GCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCT 2160
Db      1562 GCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCT 1503
Qy      2161 GGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCG 2220
Db      1502 GGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCG 1443
Qy      2221 TCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAG 2280
Db      1442 TCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAG 1383
Qy      2281 GATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTA 2340
Db      1382 GATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTA 1323
Qy      2341 CGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGG 2400
Db      1322 CGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGG 1263
Qy      2401 AAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTT 2460
Db      1262 AAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTT 1203
Qy      2461 TGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTT 2520
Db      1202 TGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTT 1143
Qy      2521 TTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAG 2580
Db      1142 TTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAG 1083
Qy      2581 ATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAAT 2640
Db      1082 ATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAAT 1023
Qy      2641 CTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACC 2700
Db      1022 CTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACC 963
Qy      2701 TATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGAT 2760
Db       962 TATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGAT 903
Qy      2761 AACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCC 2820
Db       902 AACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCC 843
Qy      2821 ACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAG 2880
Db       842 ACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAG 783
Qy      2881 AAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAG 2940
Db       782 AAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAG 723
Qy      2941 AGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGT 3000
Db       722 AGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGT 663
Qy      3001 GGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCG 3060
```

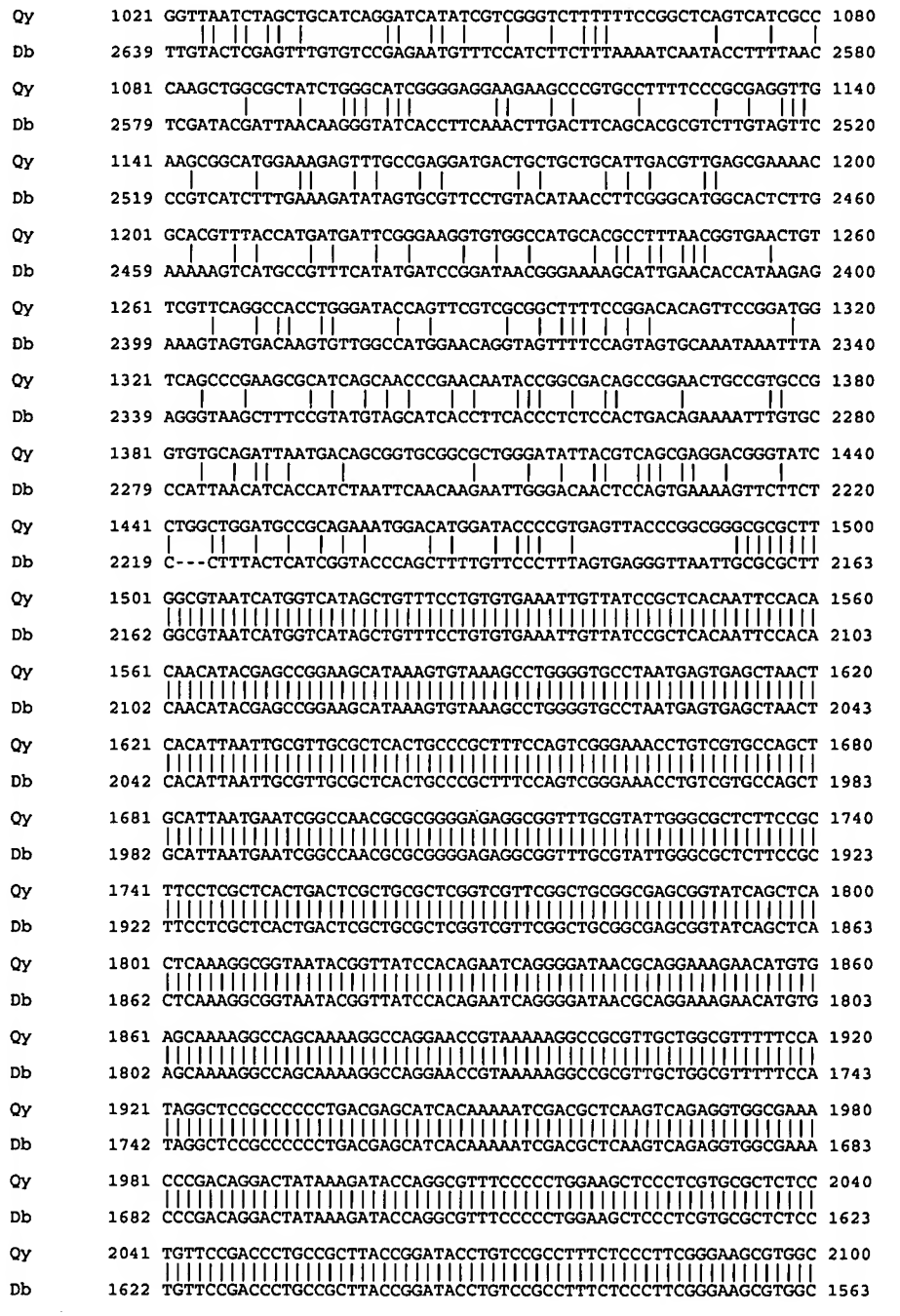
```
Db       662 GGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCG 603
Qy      3061 AGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGT 3120
Db       602 AGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGT 543
Qy      3121 TGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTC 3180
Db       542 TGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTC 483
Qy      3181 TCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTC 3240
Db       482 TCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTC 423
Qy      3241 ATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAA 3300
Db       422 ATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAA 363
Qy      3301 TACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCG 3360
Db       362 TACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCG 303
Qy      3361 AAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACC 3420
Db       302 AAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACC 243
Qy      3421 CAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAG 3480
Db       242 CAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAG 183
Qy      3481 GCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTT 3540
Db       182 GCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTT 123
Qy      3541 CCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATT 3600
Db       122 CCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATT 63
Qy      3601 TGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCC 3660
Db        62 TGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCC 3
Qy      3661 AC 3662
Db           2 AC 1
```

RESULT 12
US-10-280-913A-17/c
; Sequence 17, Application US/10280913A
; Publication No. US20040110130A1
; GENERAL INFORMATION:
; APPLICANT: Large Scale Biology Corporation
; TITLE OF INVENTION: POPULATION OF POLYNUCLEOTIDE SEQUENCE VARIANTS
; FILE REFERENCE: LSBC-PLG5392-CIP
; CURRENT APPLICATION NUMBER: US/10/280,913A
; CURRENT FILING DATE: 2002-10-25
; PRIOR APPLICATION NUMBER: 60/402,342
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 10/066,390
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: 60/268,785
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: 60/266,386
; PRIOR FILING DATE: 2002-02-02
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17
; LENGTH: 3637
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Encodes cycle 3 GFP

US-10-280-913A-17

Query Match 72.8%; Score 2664.8; DB 17; Length 3637;
Best Local Similarity 83.8%; Pred. No. 0;
Matches 3070; Conservative 0; Mismatches 567; Indels 25; Gaps 4;



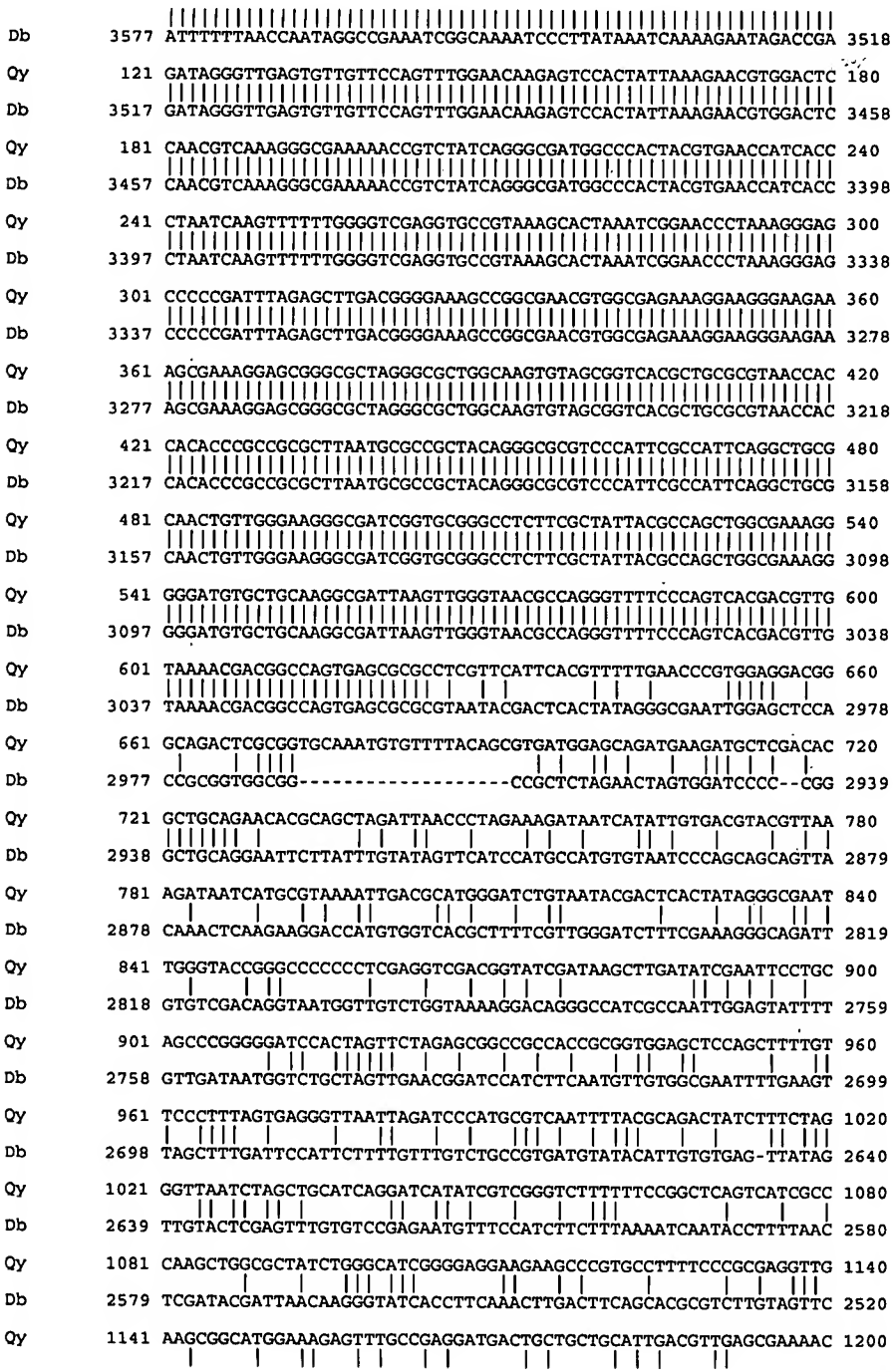


```
Qy       2101 GCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCT 2160
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1562 GCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCT 1503

Qy       2161 GGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCG 2220
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1502 GGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCG 1443

Qy       2221 TCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAG 2280
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1442 TCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAG 1383

Qy       2281 GATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTA 2340
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1382 GATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTA 1323

Qy       2341 CGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGG 2400
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1322 CGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGG 1263

Qy       2401 AAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTT 2460
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1262 AAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTT 1203

Qy       2461 TGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTT 2520
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1202 TGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTT 1143

Qy       2521 TTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAG 2580
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1142 TTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAG 1083

Qy       2581 ATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAAT 2640
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1082 ATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAAT 1023

Qy       2641 CTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACC 2700
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1022 CTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACC 963

Qy       2701 TATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGAT 2760
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        962 TATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGAT 903

Qy       2761 AACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCC 2820
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        902 AACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCC 843

Qy       2821 ACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAG 2880
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        842 ACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAG 783

Qy       2881 AAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAG 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        782 AAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAG 723

Qy       2941 AGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGT 3000
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        722 AGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGT 663

Qy       3001 GGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCG 3060
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        662 GGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCG 603

Qy       3061 AGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGT 3120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        602 AGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGT 543

Qy       3121 TGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTC 3180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        542 TGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTC 483

Qy       3181 TCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTC 3240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        482 TCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTC 423

Qy       3241 ATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAA 3300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 ATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAA 363

Qy       3301 TACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCG 3360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 TACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCG 303

Qy       3361 AAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACC 3420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 AAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACC 243

Qy       3421 CAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAG 3480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 CAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAG 183

Qy       3481 GCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTT 3540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 GCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTT 123

Qy       3541 CCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATT 3600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 CCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATT 63

Qy       3601 TGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCC 3660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          62 TGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCC 3

Qy       3661 AC 3662
             ||
Db          2 AC 1
```

RESULT 13
US-10-684-134-17/c
; Sequence 17, Application US/10684134
; Publication No. US20040142433A1
; GENERAL INFORMATION:
; APPLICANT: Large Scale Biology Corporation
; TITLE OF INVENTION: POPULATION OF POLYNUCLEOTIDE SEQUENCE VARIANTS
; FILE REFERENCE: LSBC-PLG5392-CIP
; CURRENT APPLICATION NUMBER: US/10/684,134
; CURRENT FILING DATE: 2003-10-10
; PRIOR APPLICATION NUMBER: 60/402,342
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 10/066,390
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: 60/268,785
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: 60/266,386
; PRIOR FILING DATE: 2002-02-02
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17
; LENGTH: 3637
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Encodes cycle 3 GFP
US-10-684-134-17

Query Match 72.8%; Score 2664.8; DB 17; Length 3637;
Best Local Similarity 83.8%; Pred. No. 0;
Matches 3070; Conservative 0; Mismatches 567; Indels 25; Gaps 4;

```
Qy          1 CTAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3637 CTAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTC 3578

Qy         61 ATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGA 120
```


Db 3577 ATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGA 3518
Qy 121 GATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTC 180
Db 3517 GATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTC 3458
Qy 181 CAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACC 240
Db 3457 CAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACC 3398
Qy 241 CTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAG 300
Db 3397 CTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAG 3338
Qy 301 CCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAA 360
Db 3337 CCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAA 3278
Qy 361 AGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCAC 420
Db 3277 AGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCAC 3218
Qy 421 CACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCCCATTCGCCATTCAGGCTGCG 480
Db 3217 CACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCCCATTCGCCATTCAGGCTGCG 3158
Qy 481 CAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGG 540
Db 3157 CAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGG 3098
Qy 541 GGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTG 600
Db 3097 GGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTG 3038
Qy 601 TAAAACGACGGCCAGTGAGCGCGCCTCGTTCATTCACGTTTTTGAACCCGTGGAGGACGG 660
Db 3037 TAAAACGACGGCCAGTGAGCGCGCGTAATACGACTCACTATAGGGCGAATTGGAGCTCCA 2978
Qy 661 GCAGACTCGCGGTGCAAATGTGTTTTACAGCGTGATGGAGCAGATGAAGATGCTCGACAC 720
Db 2977 CCGCGGTGGCGG----------CCGCTCTAGAACTAGTGGATCCCC--CGG 2939
Qy 721 GCTGCAGAACACGCAGCTAGATTAACCCTAGAAAGATAATCATATTGTGACGTACGTTAA 780
Db 2938 GCTGCAGGAATTCTTATTTGTATAGTTCATCCATGCCATGTGTAATCCCAGCAGCAGTTA 2879
Qy 781 AGATAATCATGCGTAAAATTGACGCATGGGATCTGTAATACGACTCACTATAGGGCGAAT 840
Db 2878 CAAACTCAAGAAGGACCATGTGGTCACGCTTTTCGTTGGGATCTTTCGAAAGGGCAGATT 2819
Qy 841 TGGGTACCGGGCCCCCCCTCGAGGTCGACGGTATCGATAAGCTTGATATCGAATTCCTGC 900
Db 2818 GTGTCGACAGGTAATGGTTGTCTGGTAAAAGGACAGGGCCATCGCCAATTGGAGTATTTT 2759
Qy 901 AGCCCGGGGGATCCACTAGTTCTAGAGCGGCCGCCACCGCGGTGGAGCTCCAGCTTTTGT 960
Db 2758 GTTGATAATGGTCTGCTAGTTGAACGGATCCATCTTCAATGTTGTGGCGAATTTTGAAGT 2699
Qy 961 TCCCTTTAGTGAGGGTTAATTAGATCCCATGCGTCAATTTTACGCAGACTATCTTTCTAG 1020
Db 2698 TAGCTTTGATTCCATTCTTTTGTTTGTCTGCCGTGATGTATACATTGTGTGAG-TTATAG 2640
Qy 1021 GGTTAATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCC 1080
Db 2639 TTGTACTCGAGTTTGTGTCCGAGAATGTTTCCATCTTCTTTAAAATCAATACCTTTTAAC 2580
Qy 1081 CAAGCTGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTG 1140
Db 2579 TCGATACGATTAACAAGGGTATCACCTTCAAACTTGACTTCAGCACGCGTCTTGTAGTTC 2520
Qy 1141 AAGCGGCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAAC 1200
Db 2519 CCGTCATCTTTGAAAGATATAGTGCGTTCCTGTACATAACCTTCGGGCATGGCACTCTTG 2460
Qy 1201 GCACGTTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGT 1260
Db 2459 AAAAAGTCATGCCGTTTCATATGATCCGGATAACGGGAAAAGCATTGAACACCATAAGAG 2400
Qy 1261 TCGTTCAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGG 1320
Db 2399 AAAGTAGTGACAAGTGTTGGCCATGGAACAGGTAGTTTTCCAGTAGTGCAAATAAATTTA 2340
Qy 1321 TCAGCCCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCG 1380
Db 2339 AGGGTAAGCTTTCCGTATGTAGCATCACCTTCACCCTCTCCACTGACAGAAAATTTGTGC 2280
Qy 1381 GTGTGCAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATC 1440
Db 2279 CCATTAACATCACCATCTAATTCAACAAGAATTGGGACAACTCCAGTGAAAAGTTCTTCT 2220
Qy 1441 CTGGCTGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCTT 1500
Db 2219 C---CTTTACTCATCGGTACCCAGCTTTTGTTCCCTTTAGTGAGGGTTAATTGCGCGCTT 2163
Qy 1501 GGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACA 1560
Db 2162 GGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACA 2103
Qy 1561 CAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACT 1620
Db 2102 CAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACT 2043
Qy 1621 CACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCT 1680
Db 2042 CACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCT 1983
Qy 1681 GCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGC 1740
Db 1982 GCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGC 1923
Qy 1741 TTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCA 1800
Db 1922 TTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCA 1863
Qy 1801 CTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTG 1860
Db 1862 CTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTG 1803
Qy 1861 AGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCA 1920
Db 1802 AGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCA 1743
Qy 1921 TAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAA 1980
Db 1742 TAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAA 1683
Qy 1981 CCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCC 2040
Db 1682 CCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCC 1623
Qy 2041 TGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGC 2100
Db 1622 TGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGC 1563
Qy 2101 GCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCT 2160
Db 1562 GCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCT 1503
Qy 2161 GGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCG 2220
Db 1502 GGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCG 1443
Qy 2221 TCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAG 2280
Db 1442 TCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAG 1383

```
Qy       2281 GATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTA 2340
Db       1382 GATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTA 1323

Qy       2341 CGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGG 2400
Db       1322 CGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGG 1263

Qy       2401 AAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTT 2460
Db       1262 AAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTT 1203

Qy       2461 TGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTT 2520
Db       1202 TGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTT 1143

Qy       2521 TTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAG 2580
Db       1142 TTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAG 1083

Qy       2581 ATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAAT 2640
Db       1082 ATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAAT 1023

Qy       2641 CTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACC 2700
Db       1022 CTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACC 963

Qy       2701 TATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGAT 2760
Db        962 TATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGAT 903

Qy       2761 AACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCC 2820
Db        902 AACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCC 843

Qy       2821 ACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAG 2880
Db        842 ACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAG 783

Qy       2881 AAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAG 2940
Db        782 AAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAG 723

Qy       2941 AGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGT 3000
Db        722 AGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGT 663

Qy       3001 GGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCG 3060
Db        662 GGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCG 603

Qy       3061 AGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGT 3120
Db        602 AGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGT 543

Qy       3121 TGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTC 3180
Db        542 TGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTC 483

Qy       3181 TCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTC 3240
Db        482 TCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTC 423

Qy       3241 ATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAA 3300
Db        422 ATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAA 363

Qy       3301 TACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCG 3360
Db        362 TACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCG 303

Qy       3361 AAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCC 3420
Db        302 AAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCC 243

Qy       3421 CAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAG 3480
Db        242 CAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAG 183

Qy       3481 GCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTT 3540
Db        182 GCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTT 123

Qy       3541 CCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATT 3600
Db        122 CCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATT 63

Qy       3601 TGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCC 3660
Db          62 TGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCC 3

Qy       3661 AC 3662
Db           2 AC 1
```

RESULT 14
US-10-637-758-17/c
; Sequence 17, Application US/10637758
; Publication No. US20040180352A1
; GENERAL INFORMATION:
; APPLICANT: Large Scale Biology Corporation
; TITLE OF INVENTION: POPULATION OF POLYNUCLEOTIDE SEQUENCE VARIANTS
; FILE REFERENCE: LSBC-PLG5392-CIP
; CURRENT APPLICATION NUMBER: US/10/637,758
; CURRENT FILING DATE: 2003-08-08
; PRIOR APPLICATION NUMBER: US/10/280,913
; PRIOR FILING DATE: 2002-10-25
; PRIOR APPLICATION NUMBER: 60/402,342
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 10/066,390
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: 60/268,785
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: 60/266,386
; PRIOR FILING DATE: 2002-02-02
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17
; LENGTH: 3637
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Encodes cycle 3 GFP
US-10-637-758-17

Query Match 72.8%; Score 2664.8; DB 17; Length 3637;
Best Local Similarity 83.8%; Pred. No. 0;
Matches 3070; Conservative 0; Mismatches 567; Indels 25; Gaps 4;

```
Qy          1 CTAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTC 60
Db       3637 CTAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTC 3578

Qy         61 ATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGA 120
Db       3577 ATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGA 3518

Qy        121 GATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTC 180
Db       3517 GATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTC 3458

Qy        181 CAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACC 240
```

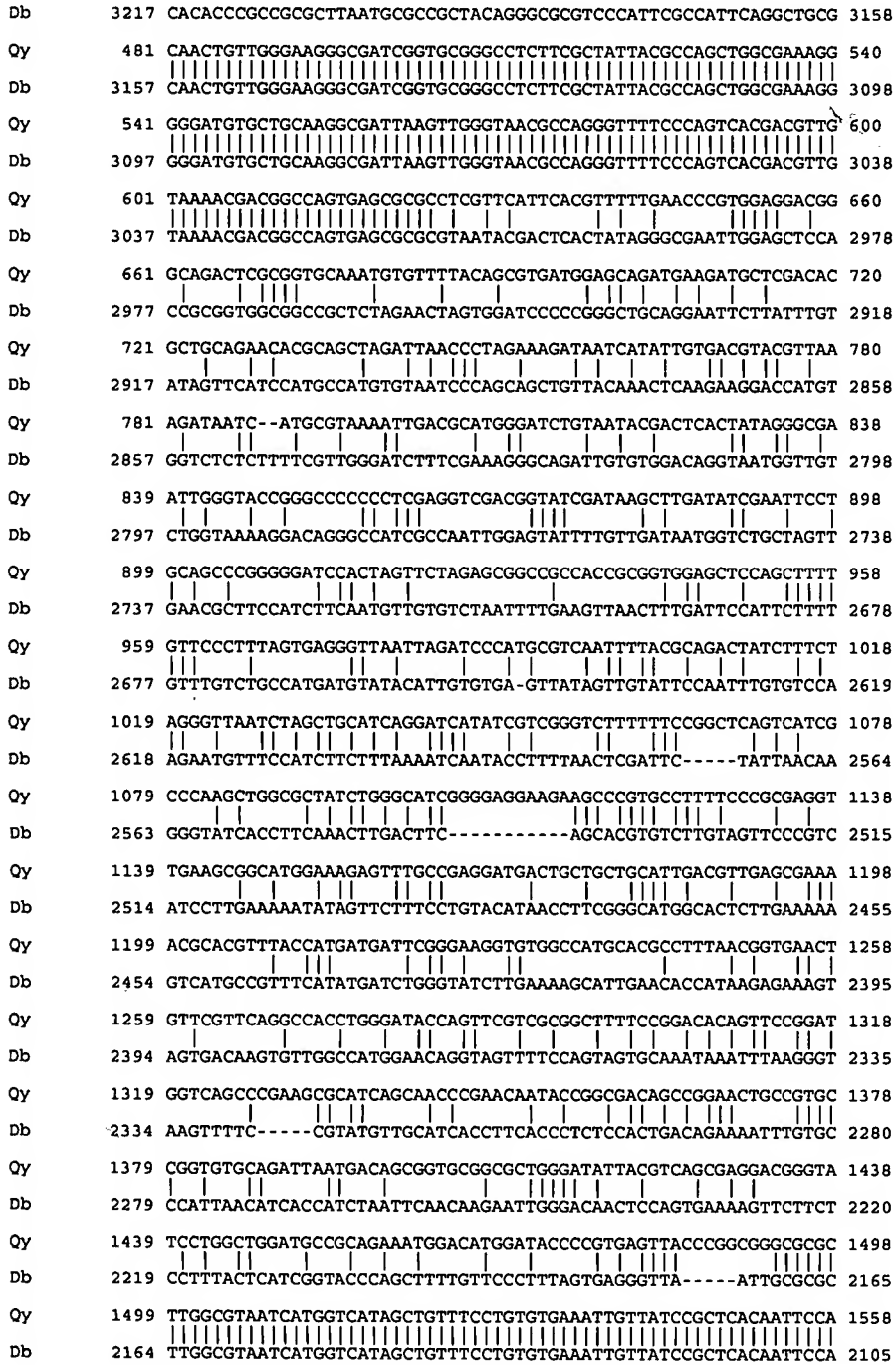
```
Db   3457 CAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACC 3398

Qy    241 CTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAG 300
Db   3397 CTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAG 3338

Qy    301 CCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAA 360
Db   3337 CCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAA 3278

Qy    361 AGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCAC 420
Db   3277 AGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCAC 3218

Qy    421 CACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCCCATTCGCCATTCAGGCTGCG 480
Db   3217 CACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCCCATTCGCCATTCAGGCTGCG 3158

Qy    481 CAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGG 540
Db   3157 CAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGG 3098

Qy    541 GGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTG 600
Db   3097 GGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTG 3038

Qy    601 TAAAACGACGGCCAGTGAGCGCGCCTCGTTCATTCACGTTTTGAACCCGTGGAGGACGG 660
Db   3037 TAAAACGACGGCCAGTGAGCGCGCGTAATACGACTCACTATAGGGCGAATTGGAGCTCCA 2978

Qy    661 GCAGACTCGCGGTGCAAATGTGTTTACAGCGTGATGGAGCAGATGAAGATGCTCGACAC 720
Db   2977 CCGCGGTGGCGG------------------CCGCTCTAGAACTAGTGGATCCCC--CGG 2939

Qy    721 GCTGCAGAACACGCAGCTAGATTAACCCTAGAAAGATAATCATATTGTGACGTACGTTAA 780
Db   2938 GCTGCAGGAATTCTTATTTGTATAGTTCATCCATGCCATGTGTAATCCCAGCAGCAGTTA 2879

Qy    781 AGATAATCATGCGTAAAATTGACGCATGGGATCTGTAATACGACTCACTATAGGGCGAAT 840
Db   2878 CAAACTCAAGAAGGACCATGTGGTCACGCTTTTCGTTGGGATCTTTCGAAAGGGCAGATT 2819

Qy    841 TGGGTACCGGGCCCCCCCTCGAGGTCGACGGTATCGATAAGCTTGATATCGAATTCCTGC 900
Db   2818 GTGTCGACAGGTAATGGTTGTCTGGTAAAAGGACAGGGCCATCGCCAATTGGAGTATTTT 2759

Qy    901 AGCCCGGGGGATCCACTAGTTCTAGAGCGGCCGCCACCGCGGTGGAGCTCCAGCTTTTGT 960
Db   2758 GTTGATAATGGTCTGCTAGTTGAACGGATCCATCTTCAATGTTGTGGCGAATTTTGAAGT 2699

Qy    961 TCCCTTTAGTGAGGGTTAATTAGATCCCATGCGTCAATTTTACGCAGACTATCTTTCTAG 1020
Db   2698 TAGCTTTGATTCCATTCTTTTGTTTGTCTGCCGTGATGTATACATTGTGTGAG-TTATAG 2640

Qy   1021 GGTTAATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCC 1080
Db   2639 TTGTACTCGAGTTTGTGTCCGAGAATGTTTCCATCTTCTTTAAAATCAATACCTTTTAAC 2580

Qy   1081 CAAGCTGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTG 1140
Db   2579 TCGATACGATTAACAAGGGTATCACCTTCAAACTTGACTTCAGCACGCGTCTTGTAGTTC 2520

Qy   1141 AAGCGGCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAAC 1200
Db   2519 CCGTCATCTTTGAAAGATATAGTGCGTTCCTGTACATAACCTTCGGGCATGGCACTCTTG 2460

Qy   1201 GCACGTTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGT 1260
Db   2459 AAAAAGTCATGCCGTTTCATATGATCCGGATAACGGGAAAAGCATTGAACACCATAAGAG 2400

Qy   1261 TCGTTCAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGG 1320
Db   2399 AAAGTAGTGACAAGTGTTGGCCATGGAACAGGTAGTTTTCCAGTAGTGCAAATAAATTTA 2340

Qy   1321 TCAGCCCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCG 1380
Db   2339 AGGGTAAGCTTTCCGTATGTAGCATCACCTTCACCCTCTCCACTGACAGAAAATTTGTGC 2280

Qy   1381 GTGTGCAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATC 1440
Db   2279 CCATTAACATCACCATCTAATTCAACAAGAATTGGGACAACTCCAGTGAAAAGTTCTTCT 2220

Qy   1441 CTGGCTGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCTT 1500
Db   2219 C---CTTTACTCATCGGTACCCAGCTTTTGTTCCCTTTAGTGAGGGTTAATTGCGCGCTT 2163

Qy   1501 GGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACA 1560
Db   2162 GGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACA 2103

Qy   1561 CAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACT 1620
Db   2102 CAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACT 2043

Qy   1621 CACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCT 1680
Db   2042 CACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCT 1983

Qy   1681 GCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGC 1740
Db   1982 GCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGC 1923

Qy   1741 TTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCA 1800
Db   1922 TTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCA 1863

Qy   1801 CTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTG 1860
Db   1862 CTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTG 1803

Qy   1861 AGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCA 1920
Db   1802 AGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCA 1743

Qy   1921 TAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAA 1980
Db   1742 TAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAA 1683

Qy   1981 CCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCC 2040
Db   1682 CCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCC 1623

Qy   2041 TGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGC 2100
Db   1622 TGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGC 1563

Qy   2101 GCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCT 2160
Db   1562 GCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCT 1503

Qy   2161 GGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCG 2220
Db   1502 GGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCG 1443

Qy   2221 TCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAG 2280
Db   1442 TCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAG 1383

Qy   2281 GATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTA 2340
Db   1382 GATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTA 1323

Qy   2341 CGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGG 2400
Db   1322 CGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGG 1263
```

```
Qy       2401 AAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTT 2460
Db       1262 AAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTT 2460

Qy       2461 TGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTT 1203
Db       1202 TGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTT 2520

Qy       2521 TTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAG 1143
Db       1142 TTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAG 2580

Qy       2581 ATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAAT 1083
Db       1082 ATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAAT 2640

Qy       2641 CTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACC 1023
Db       1022 CTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACC 2700

Qy       2701 TATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGAT 963
Db        962 TATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGAT 2760

Qy       2761 AACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCC 903
Db        902 AACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCC 2820

Qy       2821 ACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAG 843
Db        842 ACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAG 2880

Qy       2881 AAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAG 783
Db        782 AAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAG 2940

Qy       2941 AGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGT 723
Db        722 AGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGT 3000

Qy       3001 GGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCG 663
Db        662 GGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCG 3060

Qy       3061 AGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGT 603
Db        602 AGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGT 3120

Qy       3121 TGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTC 543
Db        542 TGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTC 3180

Qy       3181 TCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTC 483
Db        482 TCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTC 3240

Qy       3241 ATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAA 423
Db        422 ATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAA 3300

Qy       3301 TACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCG 363
Db        362 TACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCG 3360

Qy       3361 AAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACC 303
Db        302 AAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACC 3420

Qy       3421 CAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAG 243
Db        242 CAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAG 3480

Qy       3481 GCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTT 183
Db        182 GCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTT 3540

Qy       3541 CCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATT 123
Db         122 CCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATT 3600

Qy       3601 TGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCC 3660
Db          62 TGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCC 3

Qy       3661 AC 3662
Db           2 AC 1
```

RESULT 15
US-10-066-390-3/c
; Sequence 3, Application US/10066390
; Publication No. US20020146732A1
; GENERAL INFORMATION:
; APPLICANT: Padgett, Hal S.
; APPLICANT: Lindbo, John A.
; APPLICANT: Fitzmaurice, Wayne P.
; TITLE OF INVENTION: A Method of Increasing Complementarity
; TITLE OF INVENTION: In A Heteroduplex
; FILE REFERENCE: P-LG 4878
; CURRENT APPLICATION NUMBER: US/10/066,390
; CURRENT FILING DATE: 2002-02-01
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 3637
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic construct
US-10-066-390-3

Query Match 72.5%; Score 2656.4; DB 13; Length 3637;
Best Local Similarity 84.0%; Pred. No. 0;
Matches 3079; Conservative 0; Mismatches 556; Indels 29; Gaps 6;

```
Qy          1 CTAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTC 60
Db       3637 CTAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTC 3578

Qy         61 ATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGA 120
Db       3577 ATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGA 3518

Qy        121 GATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTC 180
Db       3517 GATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTC 3458

Qy        181 CAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACC 240
Db       3457 CAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACC 3398

Qy        241 CTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAG 300
Db       3397 CTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAG 3338

Qy        301 CCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAA 360
Db       3337 CCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAA 3278

Qy        361 AGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCAC 420
Db       3277 AGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCAC 3218

Qy        421 CACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCCCATTCGCCATTCAGGCTGCG 480
```


Db 3217 CACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCCCATTCGCCATTCAGGCTGCG 3158
Qy 481 CAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGG 540
Db 3157 CAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGG 3098
Qy 541 GGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTG 600
Db 3097 GGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTG 3038
Qy 601 TAAAACGACGGCCAGTGAGCGCGCGTAATACGACTCACTATAGGGCGAATTGGAGCTCCA 660
Db 3037 TAAAACGACGGCCAGTGAGCGCGCGTAATACGACTCACTATAGGGCGAATTGGAGCTCCA 2978
Qy 661 GCAGACTCGCGGTGCAAATGTGTTTTACAGCGTGATGGAGCAGATGAAGATGCTCGACAC 720
Db 2977 CCGCGGTGGCGGCCGCTCTAGAACTAGTGGATCCCCCGGGCTGCAGGAATTCTTATTTGT 2918
Qy 721 GCTGCAGAACACGCAGCTAGATTAACCCTAGAAAGATAATCATATTGTGACGTACGTTAA 780
Db 2917 ATAGTTCATCCATGCCATGTGTAATCCCAGCAGCTGTTACAAACTCAAGAAGGACCATGT 2858
Qy 781 AGATAATC--ATGCGTAAAATTGACGCATGGGATCTGTAATACGACTCACTATAGGGCGA 838
Db 2857 GGTCTCTCTTTTCGTTGGGATCTTTCGAAAGGGCAGATTGTGTGGACAGGTAATGGTTGT 2798
Qy 839 ATTGGGTACCGGGCCCCCCCTCGAGGTCGACGGTATCGATAAGCTTGATATCGAATTCCT 898
Db 2797 CTGGTAAAAGGACAGGGCCATCGCCAATTGGAGTATTTTGTTGATAATGGTCTGCTAGTT 2738
Qy 899 GCAGCCCGGGGGATCCACTAGTTCTAGAGCGGCCGCCACCGCGGTGGAGCTCCAGCTTTT 958
Db 2737 GAACGCTTCCATCTTCAATGTTGTGTCTAATTTTGAAGTTAACTTTGATTCCATTCTTTT 2678
Qy 959 GTTCCCTTTAGTGAGGGTTAATTAGATCCCATGCGTCAATTTTACGCAGACTATCTTTCT 1018
Db 2677 GTTTGTCTGCCATGATGTATACATTGTGTGA-GTTATAGTTGTATTCCAATTTGTGTCCA 2619
Qy 1019 AGGGTTAATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCG 1078
Db 2618 AGAATGTTTCCATCTTCTTTAAAATCAATACCTTTTAACTCGATTC-----TATTAACAA 2564
Qy 1079 CCCAAGCTGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGT 1138
Db 2563 GGGTATCACCTTCAAACTTGACTTC----------AGCACGTGTCTTGTAGTTCCCGTC 2515
Qy 1139 TGAAGCGGCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAA 1198
Db 2514 ATCCTTGAAAAATATAGTTCTTTCCTGTACATAACCTTCGGGCATGGCACTCTTGAAAAA 2455
Qy 1199 ACGCACGTTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACT 1258
Db 2454 GTCATGCCGTTTCATATGATCTGGGTATCTTGAAAAGCATTGAACACCATAAGAGAAAGT 2395
Qy 1259 GTTCGTTCAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGAT 1318
Db 2394 AGTGACAAGTGTTGGCCATGGAACAGGTAGTTTTCCAGTAGTGCAAATAAATTTAAGGGT 2335
Qy 1319 GGTCAGCCCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGC 1378
Db 2334 AAGTTTTC-----CGTATGTTGCATCACCTTCACCCTCTCCACTGACAGAAAATTTGTGC 2280
Qy 1379 CGGTGTGCAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTA 1438
Db 2279 CCATTAACATCACCATCTAATTCAACAAGAATTGGGACAACTCCAGTGAAAAGTTCTTCT 2220
Qy 1439 TCCTGGCTGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGC 1498
Db 2219 CCTTTACTCATCGGTACCCAGCTTTTGTTCCCTTTAGTGAGGGTTA-----ATTGCGCGC 2165
Qy 1499 TTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCA 1558
Db 2164 TTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCA 2105
Qy 1559 CACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAA 1618
Db 2104 CACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAA 2045
Qy 1619 CTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAG 1678
Db 2044 CTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAG 1985
Qy 1679 CTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCC 1738
Db 1984 CTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCC 1925
Qy 1739 GCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCT 1798
Db 1924 GCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCT 1865
Qy 1799 CACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATG 1858
Db 1864 CACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATG 1805
Qy 1859 TGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTC 1918
Db 1804 TGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTC 1745
Qy 1919 CATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGA 1978
Db 1744 CATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGA 1685
Qy 1979 AACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCT 2038
Db 1684 AACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCT 1625
Qy 2039 CCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTG 2098
Db 1624 CCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTG 1565
Qy 2099 GCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAG 2158
Db 1564 GCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAG 1505
Qy 2159 CTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTAT 2218
Db 1504 CTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTAT 1445
Qy 2219 CGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAAC 2278
Db 1444 CGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAAC 1385
Qy 2279 AGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAAC 2338
Db 1384 AGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAAC 1325
Qy 2339 TACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTC 2398
Db 1324 TACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTC 1265
Qy 2399 GGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTT 2458
Db 1264 GGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTT 1205
Qy 2459 TTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATC 2518
Db 1204 TTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATC 1145
Qy 2519 TTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATG 2578
Db 1144 TTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATG 1085
Qy 2579 AGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCA 2638
Db 1084 AGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCA 1025

```
Qy      2639 ATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCA 2698
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1024 ATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCA 965

Qy      2699 CCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAG 2758
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       964 CCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAG 905

Qy      2759 ATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGAC 2818
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       904 ATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGAC 845

Qy      2819 CCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGC 2878
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       844 CCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGC 785

Qy      2879 AGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCT 2938
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       784 AGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCT 725

Qy      2939 AGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATC 2998
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       724 AGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATC 665

Qy      2999 GTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGG 3058
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       664 GTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGG 605

Qy      3059 CGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATC 3118
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       604 CGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATC 545

Qy      3119 GTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAAT 3178
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       544 GTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAAT 485

Qy      3179 TCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAG 3238
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       484 TCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAG 425

Qy      3239 TCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGAT 3298
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       424 TCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGAT 365

Qy      3299 AATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGG 3358
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       364 AATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGG 305

Qy      3359 CGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCA 3418
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       304 CGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCA 245

Qy      3419 CCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGA 3478
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       244 CCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGA 185

Qy      3479 AGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTC 3538
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       184 AGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTC 125

Qy      3539 TTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATA 3598
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       124 TTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATA 65

Qy      3599 TTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTG 3658
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        64 TTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTG 5

Qy      3659 CCAC 3662
             ||||
Db         4 CCAC 1
```

Search completed: January 26, 2005, 17:51:53
Job time : 1991.66 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: January 26, 2005, 02:19:36 ; Search time 304.259 Seconds
(without alignments)
8554.922 Million cell updates/sec

Title: US-10-826-523-41
Perfect score: 3662
Sequence: 1 ctaaattgtaagcgttaata.....atttccccgaaaagtgccac 3662

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

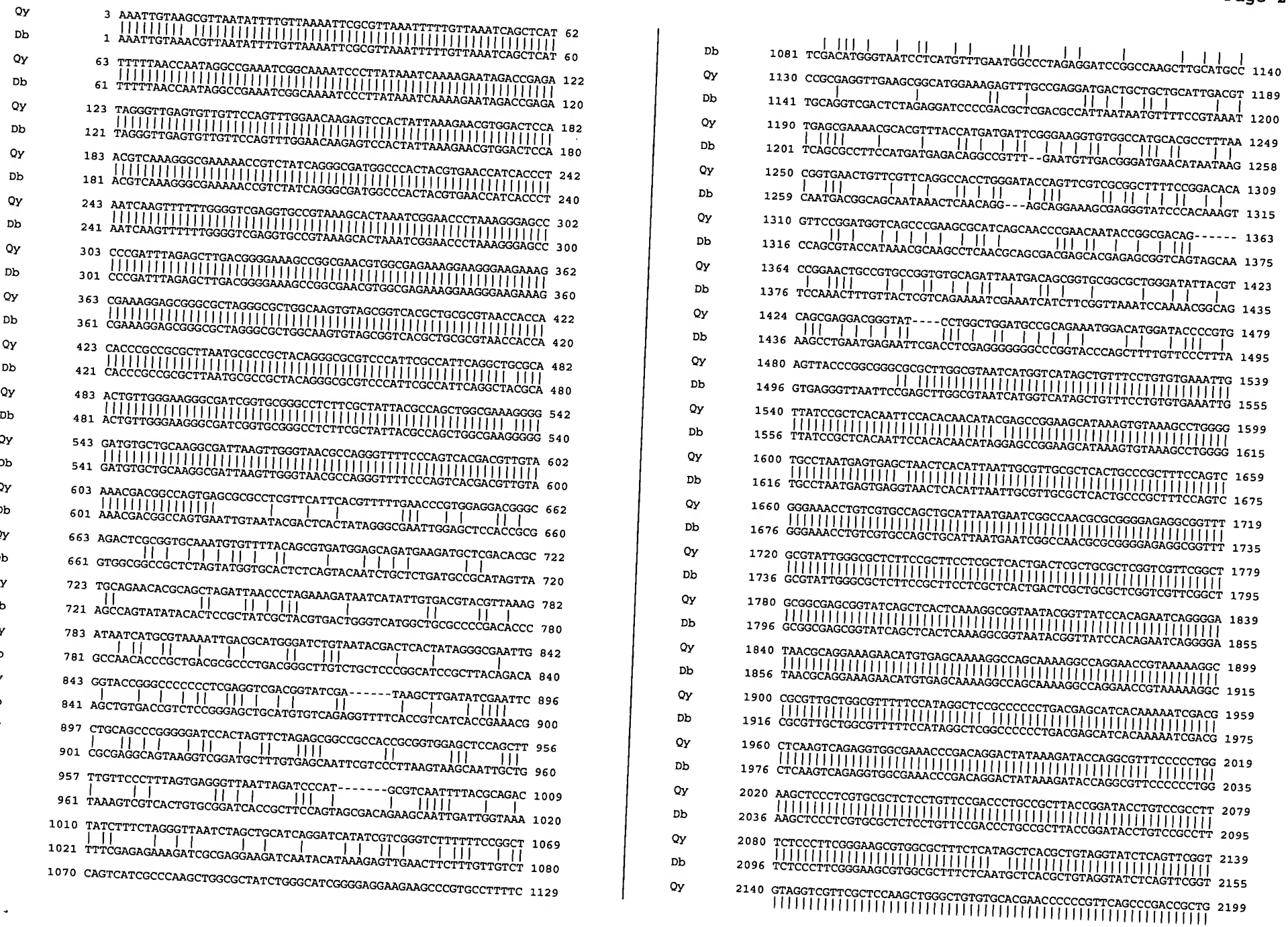
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2639	72.1	3681	1	US-07-924-028A-6	Sequence 6, Appli
2	2623.6	71.6	5314	1	US-07-924-028A-4	Sequence 4, Appli
3	2620.6	71.6	3699	3	US-08-646-538-6	Sequence 6, Appli
4	2620.6	71.6	3699	3	US-09-503-222-6	Sequence 6, Appli
5	2469.4	67.4	5277	4	US-09-890-848-16	Sequence 16, Appl
6	2448.8	66.9	3956	4	US-09-402-266B-21	Sequence 21, Appl
7	2305.2	62.9	4102	4	US-09-402-266B-20	Sequence 20, Appl
8	2302	62.9	4088	4	US-09-402-266B-18	Sequence 18, Appl
9	2279.2	62.2	4119	4	US-09-993-170-1	Sequence 1, Appli
10	2172.8	59.3	7102	3	US-09-138-024-20	Sequence 20, Appl
11	2172.8	59.3	7102	3	US-09-404-066-20	Sequence 20, Appl
12	2172.8	59.3	7102	4	US-09-573-322-20	Sequence 20, Appl
13	2172.8	59.3	7333	3	US-09-138-024-21	Sequence 21, Appl
14	2172.8	59.3	7333	3	US-09-404-066-21	Sequence 21, Appl
15	2172.8	59.3	7333	4	US-09-573-322-21	Sequence 21, Appl
16	2170.6	59.3	5178	2	US-08-474-169-2	Sequence 2, Appli
17	2170.6	59.3	6206	2	US-08-474-169-3	Sequence 3, Appli
18	2170.2	59.3	2973	4	US-09-402-266B-17	Sequence 17, Appl
19	2170.2	59.3	3789	3	US-09-075-019-8	Sequence 8, Appli
20	2170.2	59.3	4583	4	US-09-402-266B-19	Sequence 19, Appl
21	2170	59.3	2961	3	US-08-446-935-6	Sequence 6, Appli
c 22	2170	59.3	4145	3	US-08-651-472-62	Sequence 62, Appl
c 23	2170	59.3	4145	3	US-08-358-928-62	Sequence 62, Appl
c 24	2170	59.3	4277	3	US-08-651-472-63	Sequence 63, Appl
c 25	2170	59.3	4277	3	US-08-358-928-63	Sequence 63, Appl
26	2170	59.3	5356	3	US-08-446-935-1	Sequence 1, Appli
c 27	2170	59.3	5532	3	US-08-651-472-72	Sequence 72, Appl
c 28	2170	59.3	5532	3	US-08-358-928-72	Sequence 72, Appl
c 29	2170	59.3	5534	1	US-08-452-267-3	Sequence 3, Appli
c 30	2170	59.3	5534	3	US-09-123-644-3	Sequence 3, Appli
c 31	2170	59.3	6811	3	US-08-651-472-67	Sequence 67, Appl
c 32	2170	59.3	6811	3	US-08-358-928-67	Sequence 67, Appl
c 33	2170	59.3	6926	3	US-08-651-472-69	Sequence 69, Appl
c 34	2170	59.3	6926	3	US-08-358-928-69	Sequence 69, Appl
c 35	2170	59.3	7090	4	US-09-714-550-18	Sequence 18, Appl
36	2170	59.3	7214	4	US-09-438-268-1	Sequence 1, Appli
37	2170	59.3	7432	4	US-09-367-891A-6	Sequence 6, Appli
38	2170	59.3	7474	4	US-09-845-917A-26	Sequence 26, Appl
39	2170	59.3	8560	4	US-09-936-572-11	Sequence 11, Appl
40	2170	59.3	11233	3	US-08-980-832-27	Sequence 27, Appl
41	2170	59.3	11233	4	US-09-920-923B-27	Sequence 27, Appl
42	2169	59.2	3737	2	US-08-784-208-1	Sequence 1, Appli
43	2169	59.2	3796	4	US-09-470-661A-32	Sequence 32, Appl
44	2169	59.2	4399	4	US-09-911-781-27	Sequence 27, Appl
45	2169	59.2	4399	4	US-10-400-902-27	Sequence 27, Appl

ALIGNMENTS

```
RESULT 1
US-07-924-028A-6
; Sequence 6, Application US/07924028A
; Patent No. 5470573
;  GENERAL INFORMATION:
;    APPLICANT:  Lubitz Werner, Szostak, Michael P.
;    TITLE OF INVENTION:  CARRIER-BOUND RECOMBINANT PROTEINS, PROCESS
;    TITLE OF INVENTION:  FOR THE PRODUCTION AND USE AS IMMUNOGENS AND VACCINES
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Felfe & Lynch
;      STREET:  805 Third Avenue
;      CITY:  New York City
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10022
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 5.25 inch, 360 kb storage
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  PC-DOS
;      SOFTWARE:  Wordperfect
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/924,028A
;      FILING DATE:  30-SEP-1992
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/EP91/00308
;      FILING DATE:  02-FEB-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  DE 40 05 874
;      FILING DATE:  24-FEB-1990
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hanson, No. 5470573man D.
;      REGISTRATION NUMBER:  30,946
;      REFERENCE/DOCKET NUMBER:  HUBR 1027
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 688-9200
;      TELEFAX:  (212) 838-3884
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3681 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-07-924-028A-6

  Query Match             72.1%;  Score 2639;  DB 1;  Length 3681;
  Best Local Similarity   83.5%;  Pred. No. 0;
  Matches 3075;  Conservative    0;  Mismatches  580;  Indels   28;  Gaps    6;
```

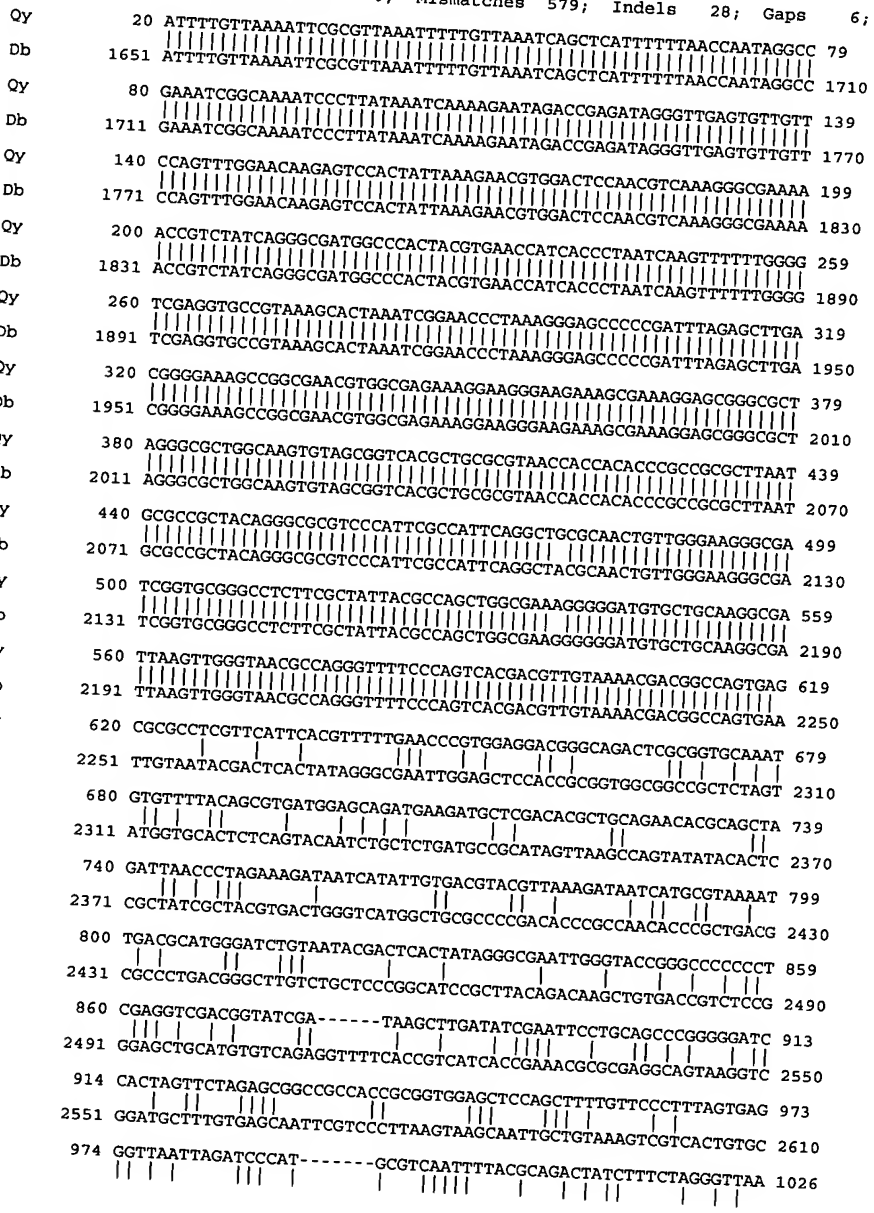


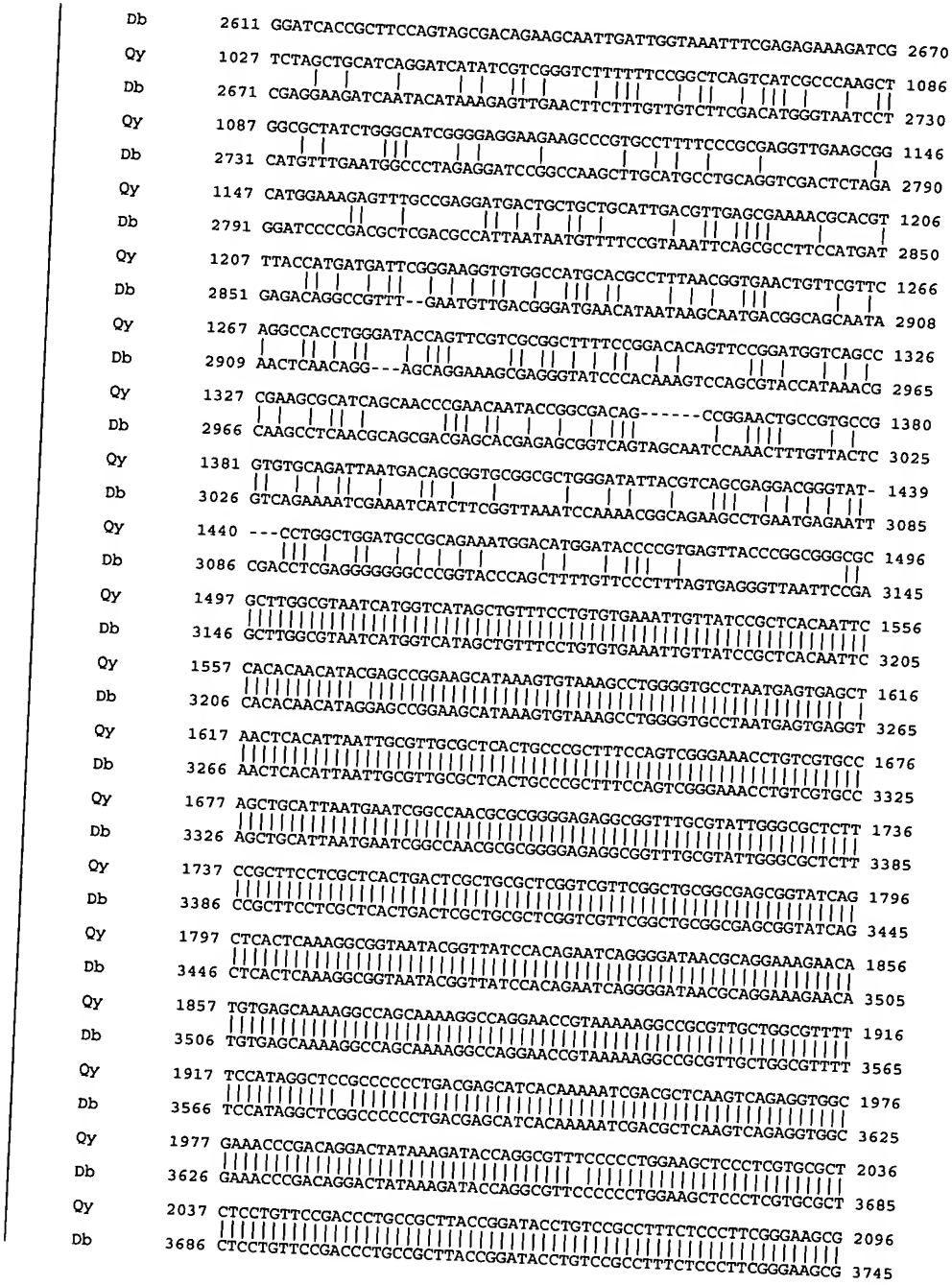
```
Db       2156 GTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTG 2215

Qy       2200 CGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACT 2259
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2216 CGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACT 2275

Qy       2260 GGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTT 2319
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2276 GGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTT 2335

Qy       2320 CTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCT 2379
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2336 CTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCT 2395

Qy       2380 GCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCAC 2439
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2396 GCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCAC 2455

Qy       2440 CGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATC 2499
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2456 CGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATC 2515

Qy       2500 TCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACG 2559
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2516 TCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACG 2575

Qy       2560 TTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTA 2619
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2576 TTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTA 2635

Qy       2620 AAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCA 2679
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2636 AAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCA 2695

Qy       2680 ATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGC 2739
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2696 ATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGC 2755

Qy       2740 CTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGC 2799
             |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2756 CTGACTGCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGC 2815

Qy       2800 TGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCC 2859
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2816 TGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCC 2875

Qy       2860 AGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTAT 2919
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2876 AGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTAT 2935

Qy       2920 TAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGT 2979
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2936 TAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGT 2995

Qy       2980 TGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTC 3039
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2996 TGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTC 3055

Qy       3040 CGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAG 3099
             |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db       3056 CGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGAAAAAAAGCGGTTAG 3115

Qy       3100 CTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGT 3159
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db       3116 CTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGCT 3175

Qy       3160 TATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGAC 3219
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3176 TATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGAC 3235

Qy       3220 TGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTG 3279
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3236 TGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTG 3295

Qy       3280 CCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCAT 3339
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3296 CCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCAT 3355

Qy       3340 TGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTC 3399
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3356 TGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTC 3415

Qy       3400 GATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTC 3459
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3416 GATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTC 3475

Qy       3460 TGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAA 3519
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3476 TGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAA 3535

Qy       3520 ATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTG 3579
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3536 ATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTG 3595

Qy       3580 TCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCG 3639
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3596 TCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCG 3655

Qy       3640 CACATTTCCCCGAAAAGTGCCAC 3662
             |||||||||||||||||||||||
Db       3656 CACATTTCCCCGAAAAGTGCCAC 3678
```

RESULT 2
US-07-924-028A-4
; Sequence 4, Application US/07924028A
; Patent No. 5470573
; GENERAL INFORMATION:
; APPLICANT: Lubitz Werner, Szostak, Michael P.
; TITLE OF INVENTION: CARRIER-BOUND RECOMBINANT PROTEINS, PROCESS
; TITLE OF INVENTION: FOR THE PRODUCTION AND USE AS IMMUNOGENS AND VACCINES
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/924,028A
; FILING DATE: 30-SEP-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP91/00308
; FILING DATE: 02-FEB-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 40 05 874
; FILING DATE: 24-FEB-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5470573man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: HUBR 1027
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5314 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single

; TOPOLOGY: linear
US-07-924-028A-4

Query Match 71.6%; Score 2623.6; DB 1; Length 5314;
Best Local Similarity 83.4%; Pred. No. 0;
Matches 3059; Conservative 0; Mismatches 579; Indels 28; Gaps 6;





```
Qy       2097 TGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCA 2156
             ||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||||||
Db       3746 TGGCGCTTTCTCAATGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCA 3805

Qy       2157 AGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACT 2216
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3806 AGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACT 3865

Qy       2217 ATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTA 2276
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3866 ATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTA 3925

Qy       2277 ACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTA 2336
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3926 ACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTA 3985

Qy       2337 ACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCT 2396
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3986 ACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCT 4045

Qy       2397 TCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTT 2456
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4046 TCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTT 4105

Qy       2457 TTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGA 2516
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4106 TTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGA 4165

Qy       2517 TCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCA 2576
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4166 TCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCA 4225

Qy       2577 TGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAAT 2636
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4226 TGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAAT 4285

Qy       2637 CAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGG 2696
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4286 CAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGG 4345

Qy       2697 CACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGT 2756
             ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db       4346 CACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTGCCCGTCGTGT 4405

Qy       2757 AGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAG 2816
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4406 AGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAG 4465

Qy       2817 ACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGC 2876
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4466 ACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGC 4525

Qy       2877 GCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAG 2936
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4526 GCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAG 4585

Qy       2937 CTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCA 2996
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4586 CTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCA 4645

Qy       2997 TCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAA 3056
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4646 TCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAA 4705

Qy       3057 GGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGA 3116
             |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db       4706 GGCGAGTTACATGATCCCCCATGTTGTGAAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGA 4765

Qy       3117 TCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATA 3176
             |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db       4766 TCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGCTTATGGCAGCACTGCATA 4825

Qy       3177 ATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCA 3236
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4826 ATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCA 4885

Qy       3237 AGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGG 3296
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4886 AGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGG 4945

Qy       3297 ATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGG 3356
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4946 ATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGG 5005

Qy       3357 GGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTG 3416
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5006 GGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTG 5065

Qy       3417 CACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAG 3476
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5066 CACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAG 5125

Qy       3477 GAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATAC 3536
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5126 GAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATAC 5185

Qy       3537 TCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACA 3596
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5186 TCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACA 5245

Qy       3597 TATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAG 3656
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5246 TATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAG 5305

Qy       3657 TGCCAC 3662
             ||||||
Db       5306 TGCCAC 5311
```

RESULT 3
US-08-646-538-6
; Sequence 6, Application US/08646538
; Patent No. 6027881
; GENERAL INFORMATION:
; APPLICANT: Pavlakis, George N.
; APPLICANT: Gaitanaris, George A.
; APPLICANT: Stauber, Roland H.
; APPLICANT: Vournakis, John N.
; TITLE OF INVENTION: Mutant Aequorea victoria Fluorescent
; TITLE OF INVENTION: Proteins Having Increased Cellular Fluorescence
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/646,538
; FILING DATE: No. 6027881 yet assigned
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Kenneth A.
; REGISTRATION NUMBER: 31,677
; REFERENCE/DOCKET NUMBER: 015280-249000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 6:

```
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3699 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  1..3699
;      OTHER INFORMATION:  /note= "pBSGFP"
US-08-646-538-6

  Query Match             71.6%;  Score 2620.6;  DB 3;  Length 3699;
  Best Local Similarity   83.3%;  Pred. No. 0;
  Matches 3078;  Conservative    0;  Mismatches  579;  Indels   40;  Gaps    7;

Qy          3 AAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTCAT 62
              |||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3 AAATTGTAAACGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTCAT 62

Qy         63 TTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGAGA 122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 TTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGAGA 122

Qy        123 TAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTCCA 182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 TAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTCCA 182

Qy        183 ACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCCT 242
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 ACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCCT 242

Qy        243 AATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAGCC 302
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 AATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAGCC 302

Qy        303 CCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAAAG 362
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 CCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAAAG 362

Qy        363 CGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCACCA 422
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 CGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCACCA 422

Qy        423 CACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTC---CCATTCGCCATTCAGGCTGC 479
              |||||||||||||||||||||||||||||||||||||   |||||||||||||||||||
Db        423 CACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCGCGCCATTCGCCATTCAGGCTGC 482

Qy        480 GCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAG 539
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        483 GCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAG 542

Qy        540 GGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTT 599
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        543 GGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTT 602

Qy        600 GTAAAACGACGGCCAGTGAGCGCGCCTCGTTCATTCACGTTTTTGAACCCGTGGAGGACG 659
              |||||||||||||||||||  |  | |||| |  |  |  |  ||  |  |  |  |||
Db        603 GTAAAACGACGGCCAGTGAATTGTAAT--ACGACTCACTATAGGGCGAATTGGGTACCGG 660

Qy        660 GGCAGACTCGCGGTGCAAATGTGTTTTACAGCGTGATGGAGCAGATGAAGATGCTCGACA 719
              | | |  |  | | || |  ||| |||| |  ||  ||  | |
Db        661 GCCCCCCCTCGAGGTCGACGGTATCGATAAGCTTGATGATCCTTATTTGTATAGTTCATC 720

Qy        720 CGCTGCA---GAACACGCAGCTAGATTAACCCTAGAAAGATAATCATATTGTGACGTACG 776
              |  ||  |  | | ||||   | ||  |  ||||  |  ||  |  |  |  |  |
Db        721 CATGCCATGTGTAATCCCAGCAGCTGTTACAAACTCAAGAAGGACCATGTGGTCTCTCTT 780

Qy        777 TTAAAGATAATCATGCGTAAAATTGACGCATGGGATCTGTAATACGACTCACTATAGGGC 836
              ||  ||| | || ||           |||  | | | | | | | ||
Db        781 TTCGTTGGGATCTTTCGAAAGGGCAGATTGTGTGGACAGGTA-ATGGTTGTCTGGTAAAA 839

Qy        837 GAATTGGGTACCGGGCCCCCCCTCGAGGTCGACGGTATCGATAAGCTTGATATCGAATTC 896
```

```
Db        840 GGACAGGGCCATCGCCAATTGGAGTATTTTGTTGATAATGGTCTGCTAGTTGAACGCTTC 899

Qy        897 CTGCAGCCCGGGGGATCCACTAGTT----CTAGAGCGGCCGCCACCGCGGTGGAGCTCCA 952
              |  |  |  |  |  |  ||  |    |  ||  |  |  |  |||  ||  ||
Db        900 CATCTTCAATGTTGTGTCTAATTTTGAAGTTAACTTTGATTCCATTCTTTTGTTTGTCTG 959

Qy        953 GCTTTTGTTCCCTTTAGTGAGGGTTAATTAGATCCCATGCGTCAATTTTACGCAGACTAT 1012
              |||  |  |  |  |||| |||| |  |  |  |  |  |  |  |  |  |  |
Db        960 CCATGATGTATACATTGTGTGAGTTATAGTTGTATTCCAATTTGTGTCCAAGAATGTTTC 1019

Qy       1013 CTTTCTAGGGTTAATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAG 1072
              |||  |  |||| |  |  |  |  |  |  |  |  |  |||  |||
Db       1020 CATCTTCTTTAAAATCAATACCTTTTAACTCGATTCTATTAACAAGGGTATCACCTTCAA 1079

Qy       1073 TCATCGCCCAAGCTGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCG 1132
              |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db       1080 ACTTGACTTCAGCACGTGTCTTGTAGTTCCCGTCATCTTTGAAAAATATAGTTCTTTCCT 1139

Qy       1133 CGAGGTTGAAGCGGCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGA 1192
              |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db       1140 GTACATAACCTTCGGGCATGGCACTCTTGAAAAAGTCATGCCGTTTCATATGATCCGGGT 1199

Qy       1193 GCGAAAACGCACGTTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGG 1252
              |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db       1200 ATCTTGAAAAGCATTGAACACCATAAGAGAAAGTAGTGACAAGTGTTGGCCATGGAACAG 1259

Qy       1253 TGAACTGTTCGTTCAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTT 1312
              |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db       1260 GTAGTTTTCCAGTAGTGCAAATAAATTTAAGGGTAAGTTTTCCGTATGTTGCATCACCTT 1319

Qy       1313 CCGGATGGTCAGCCCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTG 1372
              |  |  ||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db       1320 CACCCTCTCCACTGACAGAAAATTTGTGCCCATTAACATCACCATCTAATTCAACAAGAA 1379

Qy       1373 CCGTGCC----------GGTGTGCAGATTAATGACAGCGGTGCGGCGCTGGGATAT 1418
              |  |  |          |  |  |  |  |  |  |  |  |  |  |  ||
Db       1380 TTGGGACAACTCCAGTGAAGAGTTCTTCTCCTTTGCTAGCCATTTCTTGCGCGATCGAAT 1439

Qy       1419 TACGTCAGCGAGGACGGGTATCCTGGCTGGATGCCGCAGAAATGGACATGGATACCCCGT 1478
              |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db       1440 TCCTGCAGCCCGGGGGATCCACTAGTTCTAGAGCGGCCGCCACCGCGGTGGAGCTCCAGC 1499

Qy       1479 GAGTTACCCGGCGGG----------CGCGCTTGGCGTAATCATGGTCATAGCTGTTT 1525
              |  |  ||  |  |          |  |  |  |  |  |  |  |  |  |  |
Db       1500 TTTTGTTCCCTTTAGTGAGGGTTAATTCCGAGCTTGGCGTAATCATGGTCATAGCTGTTT 1559

Qy       1526 CCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAG 1585
              |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db       1560 CCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAG 1619

Qy       1586 TGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTG 1645
              |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db       1620 TGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTG 1679

Qy       1646 CCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCG 1705
              |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db       1680 CCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCG 1739

Qy       1706 GGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGC 1765
              |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db       1740 GGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGC 1799

Qy       1766 TCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCC 1825
              |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db       1800 TCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCC 1859

Qy       1826 ACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGG 1885
              |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db       1860 ACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGG 1919

Qy       1886 AACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCAT 1945
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1920 AACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCAT 1979

Qy       1946 CACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAG 2005
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1980 CACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAG 2039

Qy       2006 GCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGA 2065
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2040 GCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGA 2099

Qy       2066 TACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGG 2125
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2100 TACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGG 2159

Qy       2126 TATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTT 2185
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2160 TATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTT 2219

Qy       2186 CAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACAC 2245
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2220 CAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACAC 2279

Qy       2246 GACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGC 2305
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2280 GACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGC 2339

Qy       2306 GGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTT 2365
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2340 GGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTT 2399

Qy       2366 GGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCC 2425
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2400 GGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCC 2459

Qy       2426 GGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGC 2485
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2460 GGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGC 2519

Qy       2486 AGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGG 2545
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2520 AGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGG 2579

Qy       2546 AACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAG 2605
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2580 AACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAG 2639

Qy       2606 ATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGG 2665
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2640 ATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGG 2699

Qy       2666 TCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGT 2725
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2700 TCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGT 2759

Qy       2726 TCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCA 2785
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2760 TCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCA 2819

Qy       2786 TCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCA 2845
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2820 TCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCA 2879

Qy       2846 GCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCC 2905
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2880 GCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCC 2939

Qy       2906 TCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGT 2965
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2940 TCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGT 2999

Qy       2966 TTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATG 3025
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3000 TTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATG 3059

Qy       3026 GCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGC 3085
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3060 GCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGC 3119

Qy       3086 AAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTG 3145
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3120 AAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTG 3179

Qy       3146 TTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGA 3205
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3180 TTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGA 3239

Qy       3206 TGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGA 3265
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3240 TGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGA 3299

Qy       3266 CCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTA 3325
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3300 CCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTA 3359

Qy       3326 AAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTG 3385
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3360 AAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTG 3419

Qy       3386 TTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACT 3445
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3420 TTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACT 3479

Qy       3446 TTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATA 3505
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3480 TTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATA 3539

Qy       3506 AGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATT 3565
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3540 AGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATT 3599

Qy       3566 TATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAA 3625
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3600 TATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAA 3659

Qy       3626 ATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCAC 3662
              |||||||||||||||||||||||||||||||||||||
Db       3660 ATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCAC 3696
```

```
RESULT 4
US-09-503-222-6
; Sequence 6, Application US/09503222
; Patent No. 6265548
;  GENERAL INFORMATION:
;    APPLICANT:  Pavlakis, George N.
;    APPLICANT:  Gaitanaris, George A.
;    APPLICANT:  Stauber, Roland H.
;    APPLICANT:  Vournakis, John N.
;    TITLE OF INVENTION:  Mutant Aequorea victoria Fluorescent
;    TITLE OF INVENTION:  Proteins Having Increased Cellular Fluorescence
;    NUMBER OF SEQUENCES:  37
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, 8th Floor
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/503,222
;      FILING DATE:
```

; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/646,538
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Kenneth A.
; REGISTRATION NUMBER: 31,677
; REFERENCE/DOCKET NUMBER: 015280-249000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3699 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; FEATURE:
; NAME/KEY: -
; LOCATION: 1..3699
; OTHER INFORMATION: /note= "pBSGFP"
US-09-503-222-6

Query Match 71.6%; Score 2620.6; DB 3; Length 3699;
Best Local Similarity 83.3%; Pred. No. 0;
Matches 3078; Conservative 0; Mismatches 579; Indels 40; Gaps 7;

```
Qy        3 AAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTCAT 62
Db        3 AAATTGTAAACGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTCAT 62

Qy       63 TTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGAGA 122
Db       63 TTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGAGA 122

Qy      123 TAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTCCA 182
Db      123 TAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTCCA 182

Qy      183 ACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCCT 242
Db      183 ACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCCT 242

Qy      243 AATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAGCC 302
Db      243 AATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAGCC 302

Qy      303 CCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAAAG 362
Db      303 CCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAAAG 362

Qy      363 CGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCACCA 422
Db      363 CGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCACCA 422

Qy      423 CACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTC---CCATTCGCCATTCAGGCTGC 479
Db      423 CACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCGCGCCATTCGCCATTCAGGCTGC 482

Qy      480 GCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAG 539
Db      483 GCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAG 542

Qy      540 GGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTT 599
Db      543 GGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTT 599

Qy      600 GTAAAACGACGGCCAGTGAGCGCGCCTCGTTCATTCACGTTTTGAACCCGTGGAGGACG 659
Db      603 GTAAAACGACGGCCAGTGAATTGTAAT--ACGACTCACTATAGGGCGAATTGGGTACCGG 660

Qy      660 GGCAGACTCGCGGTGCAAATGTGTTTTACAGCGTGATGGAGCAGATGAAGATGCTCGACA 719
```

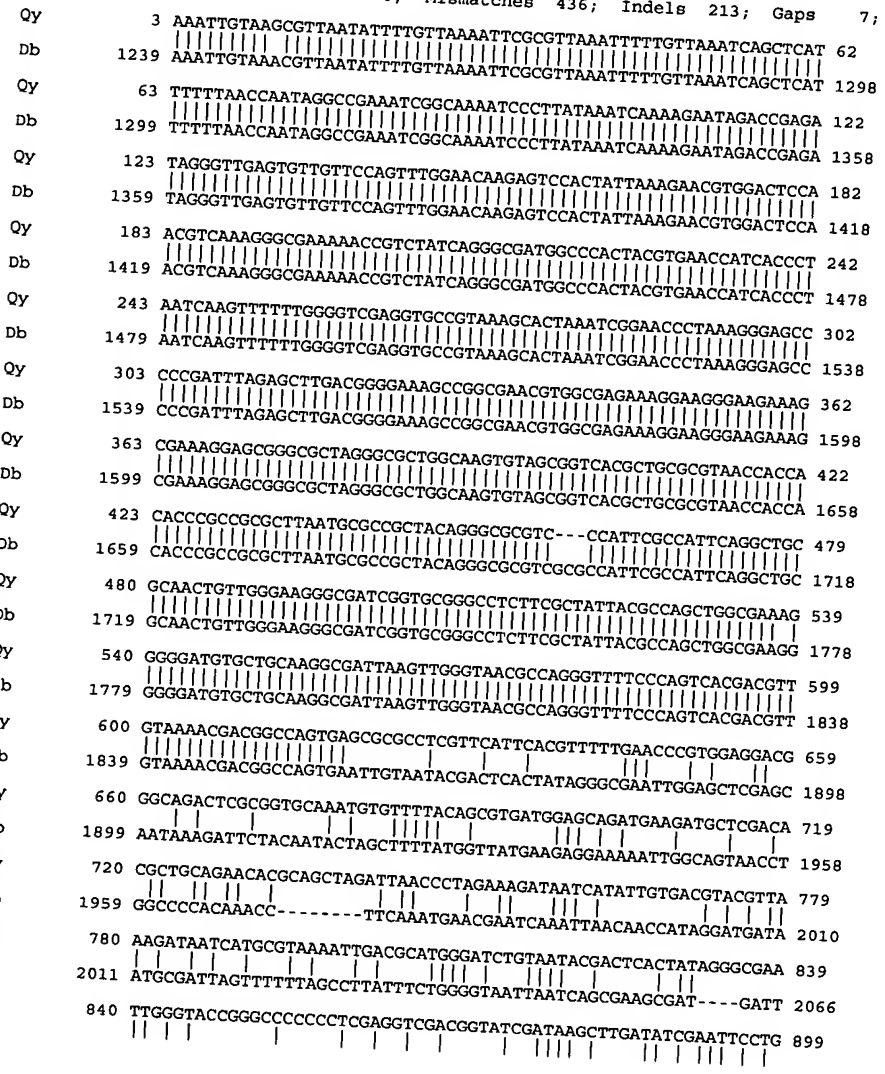
```
Db      661 GCCCCCCCTCGAGGTCGACGGTATCGATAAGCTTGATGATCCTTATTTGTATAGTTCATC 720

Qy      720 CGCTGCA---GAACACGCAGCTAGATTAACCCTAGAAAGATAATCATATTGTGACGTACG 776
Db      721 CATGCCATGTGTAATCCCAGCAGCTGTTACAAACTCAAGAAGGACCATGTGGTCTCTCTT 780

Qy      777 TTAAAGATAATCATGCGTAAAATTGACGCATGGGATCTGTAATACGACTCACTATAGGGC 836
Db      781 TTCGTTGGGATCTTTCGAAAGGGCAGATTGTGTGGACAGGTA-ATGGTTGTCTGGTAAAA 839

Qy      837 GAATTGGGTACCGGGCCCCCCCTCGAGGTCGACGGTATCGATAAGCTTGATATCGAATTC 896
Db      840 GGACAGGGCCATCGCCAATTGGAGTATTTTGTTGATAATGGTCTGCTAGTTGAACGCTTC 899

Qy      897 CTGCAGCCCGGGGGATCCACTAGTT----CTAGAGCGGCCGCCACCGCGGTGGAGCTCCA 952
Db      900 CATCTTCAATGTTGTGTCTAATTTTGAAGTTAACTTTGATTCCATTCTTTTGTTTGTCTG 959

Qy      953 GCTTTTGTTCCCTTTAGTGAGGGTTAATTAGATCCCATGCGTCAATTTTACGCAGACTAT 1012
Db      960 CCATGATGTATACATTGTGTGAGTTATAGTTGTATTCCAATTTGTGTCCAAGAATGTTTC 1019

Qy     1013 CTTTCTAGGGTTAATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAG 1072
Db     1020 CATCTTCTTTAAAATCAATACCTTTAACTCGATTCTATTAACAAGGGTATCACCTTCAA 1079

Qy     1073 TCATCGCCCAAGCTGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCG 1132
Db     1080 ACTTGACTTCAGCACGTGTCTTGTAGTTCCCGTCATCTTTGAAAAATATAGTTCTTTCCT 1139

Qy     1133 CGAGGTTGAAGCGGCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGA 1192
Db     1140 GTACATAACCTTCGGGCATGGCACTCTTGAAAAAGTCATGCCGTTTCATATGATCCGGGT 1199

Qy     1193 GCGAAAACGCACGTTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGG 1252
Db     1200 ATCTTGAAAAGCATTGAACACCATAAGAGAAAGTAGTGACAAGTGTTGGCCATGGAACAG 1259

Qy     1253 TGAACTGTTCGTTCAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTT 1312
Db     1260 GTAGTTTTCCAGTAGTGCAAATAAATTTAAGGGTAAGTTTTCCGTATGTTGCATCACCTT 1319

Qy     1313 CCGGATGGTCAGCCCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTG 1372
Db     1320 CACCCTCTCCACTGACAGAAAATTTGTGCCCATTAACATCACCATCTAATTCAACAAGAA 1379

Qy     1373 CCGTGCC----------GGTGTGCAGATTAATGACAGCGGTGCGGCGCTGGGATAT 1418
Db     1380 TTGGGACAACTCCAGTGAAGAGTTCTTCTCCTTTGCTAGCCATTTCTTGCGCGATCGAAT 1439

Qy     1419 TACGTCAGCGAGGACGGGTATCCTGGCTGGATGCCGCAGAAATGGACATGGATACCCCGT 1478
Db     1440 TCCTGCAGCCCGGGGGATCCACTAGTTCTAGAGCGGCCGCCACCGCGGTGGAGCTCCAGC 1499

Qy     1479 GAGTTACCCGGCGGG----------CGCGCTTGGCGTAATCATGGTCATAGCTGTTT 1525
Db     1500 TTTTGTTCCCTTTAGTGAGGGTTAATTCCGAGCTTGGCGTAATCATGGTCATAGCTGTTT 1559

Qy     1526 CCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAG 1585
Db     1560 CCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAG 1619

Qy     1586 TGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTG 1645
Db     1620 TGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTG 1679

Qy     1646 CCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCG 1705
Db     1680 CCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCG 1739

Qy     1706 GGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGC 1765
```

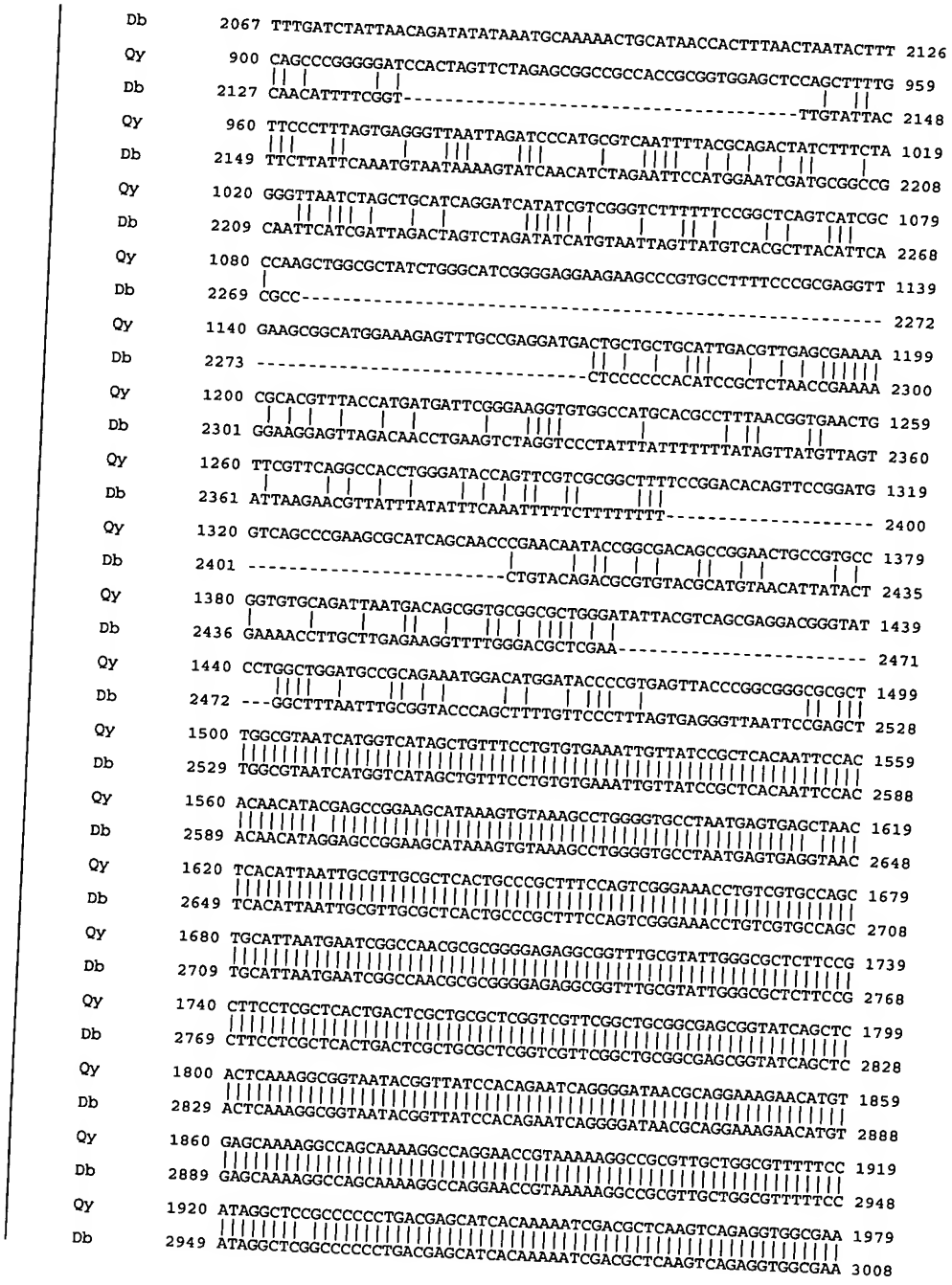
```
Db       1740 GGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGC 1799
Qy       1766 TCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCC 1825
Db       1800 TCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCC 1859
Qy       1826 ACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGG 1885
Db       1860 ACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGG 1919
Qy       1886 AACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCAT 1945
Db       1920 AACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCAT 1979
Qy       1946 CACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAG 2005
Db       1980 CACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAG 2039
Qy       2006 GCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGA 2065
Db       2040 GCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGA 2099
Qy       2066 TACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGG 2125
Db       2100 TACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGG 2159
Qy       2126 TATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTT 2185
Db       2160 TATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTT 2219
Qy       2186 CAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACAC 2245
Db       2220 CAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACAC 2279
Qy       2246 GACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGC 2305
Db       2280 GACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGC 2339
Qy       2306 GGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTT 2365
Db       2340 GGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTT 2399
Qy       2366 GGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCC 2425
Db       2400 GGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCC 2459
Qy       2426 GGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGC 2485
Db       2460 GGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGC 2519
Qy       2486 AGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGG 2545
Db       2520 AGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGG 2579
Qy       2546 AACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAG 2605
Db       2580 AACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAG 2639
Qy       2606 ATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGG 2665
Db       2640 ATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGG 2699
Qy       2666 TCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGT 2725
Db       2700 TCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGT 2759
Qy       2726 TCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCA 2785
Db       2760 TCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCA 2819
Qy       2786 TCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCA 2845
Db       2820 TCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCA 2879

Qy       2846 GCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCC 2905
Db       2880 GCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCC 2939
Qy       2906 TCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGT 2965
Db       2940 TCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGT 2999
Qy       2966 TTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATG 3025
Db       3000 TTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATG 3059
Qy       3026 GCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGC 3085
Db       3060 GCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGC 3119
Qy       3086 AAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTG 3145
Db       3120 AAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTG 3179
Qy       3146 TTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGA 3205
Db       3180 TTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGA 3239
Qy       3206 TGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGA 3265
Db       3240 TGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGA 3299
Qy       3266 CCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTA 3325
Db       3300 CCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTA 3359
Qy       3326 AAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTG 3385
Db       3360 AAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTG 3419
Qy       3386 TTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACT 3445
Db       3420 TTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACT 3479
Qy       3446 TTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATA 3505
Db       3480 TTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATA 3539
Qy       3506 AGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATT 3565
Db       3540 AGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATT 3599
Qy       3566 TATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAA 3625
Db       3600 TATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAA 3659
Qy       3626 ATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCAC 3662
Db       3660 ATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCAC 3696
```

RESULT 5
US-09-890-848-16
; Sequence 16, Application US/09890848
; Patent No. 6573053
; GENERAL INFORMATION:
; APPLICANT: Firth, Greg
; APPLICANT: Odedra, Rajesh Muru
; TITLE OF INVENTION: Analysis Method
; FILE REFERENCE: PA9903
; CURRENT APPLICATION NUMBER: US/09/890,848
; CURRENT FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: PCT/GB00/00346
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: EP 99300873.9
; PRIOR FILING DATE: 1999-02-05

; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
; LENGTH: 5277
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: plasmid
US-09-890-848-16

Query Match 67.4%; Score 2469.4; DB 4; Length 5277;
Best Local Similarity 82.3%; Pred. No. 0;
Matches 3014; Conservative 0; Mismatches 436; Indels 213; Gaps 7;



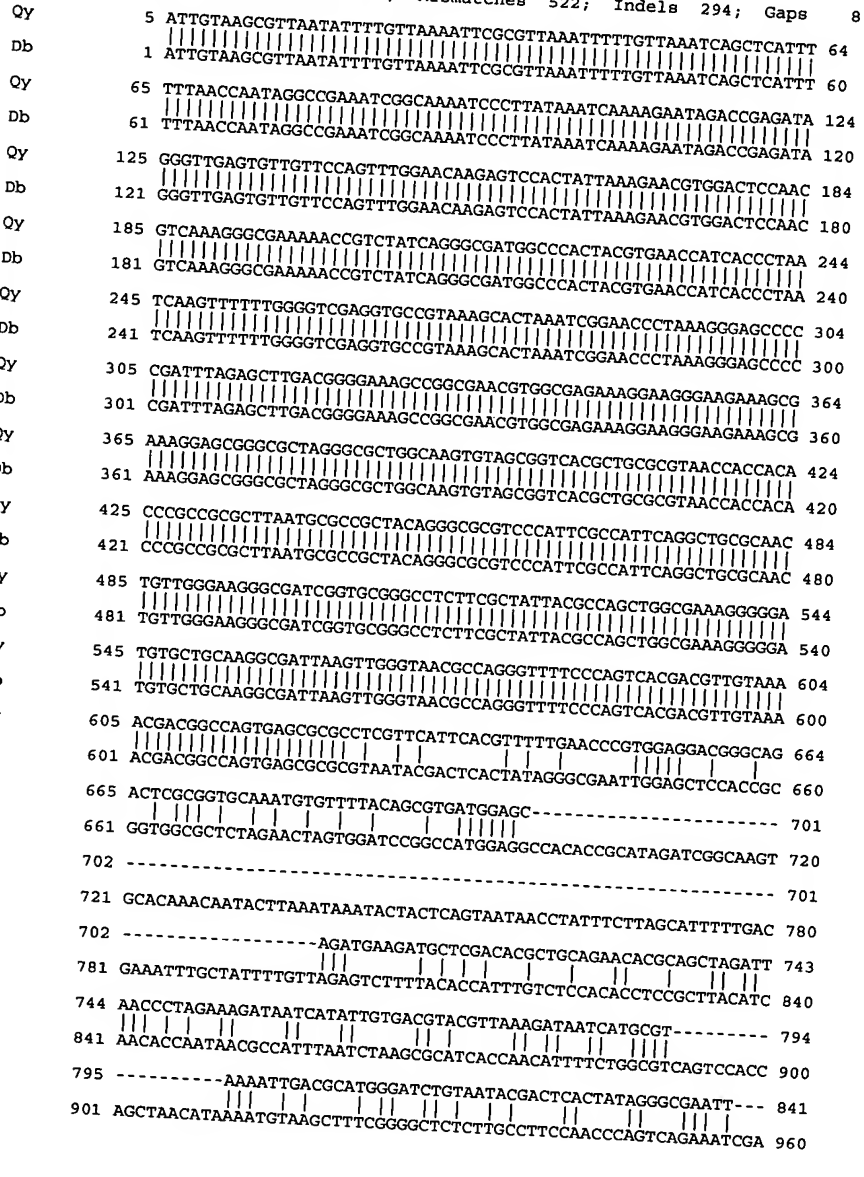


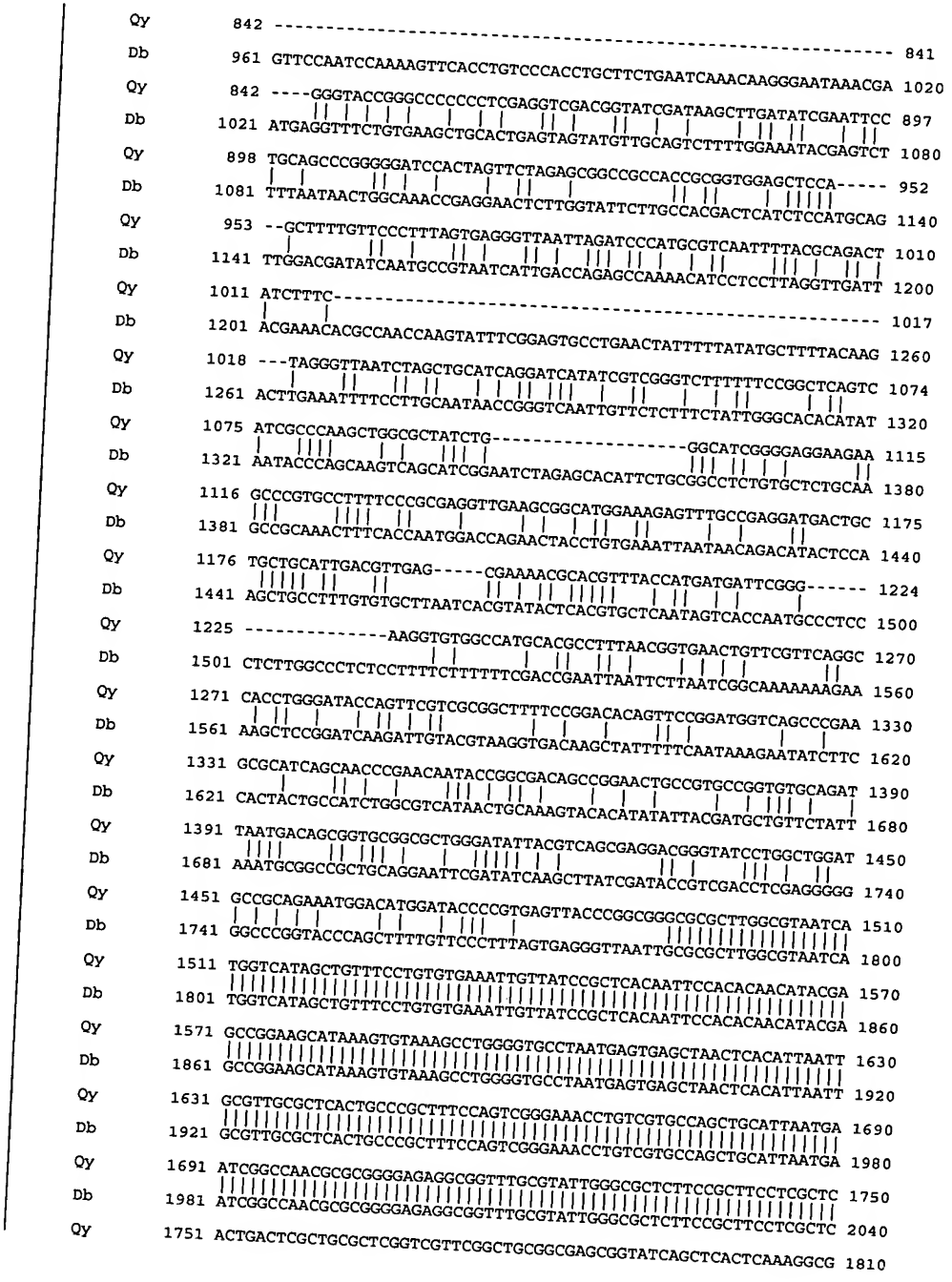
```
Qy       1980 ACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTC 2039
             ||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db       3009 ACCCGACAGGACTATAAAGATACCAGGCGTTCCCCCCTGGAAGCTCCCTCGTGCGCTCTC 3068

Qy       2040 CTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGG 2099
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3069 CTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGG 3128

Qy       2100 CGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGC 2159
             |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3129 CGCTTTCTCAATGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGC 3188

Qy       2160 TGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATC 2219
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3189 TGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATC 3248

Qy       2220 GTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACA 2279
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3249 GTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACA 3308

Qy       2280 GGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACT 2339
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3309 GGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACT 3368

Qy       2340 ACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCG 2399
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3369 ACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCG 3428

Qy       2400 GAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTT 2459
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3429 GAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTT 3488

Qy       2460 TTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCT 2519
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3489 TTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCT 3548

Qy       2520 TTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGA 2579
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3549 TTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGA 3608

Qy       2580 GATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAA 2639
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3609 GATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAA 3668

Qy       2640 TCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCAC 2699
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3669 TCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCAC 3728

Qy       2700 CTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGA 2759
             |||||||||||||||||||||||||||||||||||||||||||||| |||||  |||||||
Db       3729 CTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTGCCCGTGTGGTAGA 3788

Qy       2760 TAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACC 2819
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3789 TAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACC 3848

Qy       2820 CACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCA 2879
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3849 CACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCA 3908

Qy       2880 GAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTA 2939
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3909 GAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTA 3968

Qy       2940 GAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCG 2999
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3969 GAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCG 4028

Qy       3000 TGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGC 3059
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4029 TGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGC 4088

Qy       3060 GAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCG 3119
             ||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db       4089 GAGTTACATGATCCCCCATGTTGTGAAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCG 4148

Qy       3120 TTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATT 3179
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4149 TTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATT 4208

Qy       3180 CTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGT 3239
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4209 CTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGT 4268

Qy       3240 CATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATA 3299
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4269 CATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATA 4328

Qy       3300 ATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGC 3359
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4329 ATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGC 4388

Qy       3360 GAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCAC 3419
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4389 GAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCAC 4448

Qy       3420 CCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAA 3479
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4449 CCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAA 4508

Qy       3480 GGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCT 3539
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4509 GGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCT 4568

Qy       3540 TCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATAT 3599
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4569 TCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATAT 4628

Qy       3600 TTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGC 3659
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4629 TTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGC 4688

Qy       3660 CAC 3662
             |||
Db       4689 CAC 4691
```

RESULT 6
US-09-402-266B-21
; Sequence 21, Application US/09402266B
; Patent No. 6537767
; GENERAL INFORMATION:
; APPLICANT: HINNEN, Albert
; APPLICANT: HEGEMANN, Johannes
; APPLICANT: MUNDER, Thomas
; APPLICANT: SCHUSTER, Tilmer
; APPLICANT: FELDMANN, Horst
; APPLICANT: KRAMER, Wilfried
; APPLICANT: ZIMMERMANN, Friedrich
; APPLICANT: ENTIAN, Karl-Dieter
; TITLE OF INVENTION: Process for Screening Antimycotically Active Substances
; FILE REFERENCE: 38005-0094
; CURRENT APPLICATION NUMBER: US/09/402,266B
; CURRENT FILING DATE: 1999-10-08
; PRIOR APPLICATION NUMBER: PCT/EP98/01904
; PRIOR FILING DATE: 1998-10-08
; PRIOR APPLICATION NUMBER: DE 19713572.2
; PRIOR FILING DATE: 1997-04-02
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 21
; LENGTH: 3956
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:

; NAME/KEY: misc_feature
; LOCATION: ()..()
; OTHER INFORMATION: Synthetic sequence
US-09-402-266B-21

Query Match 66.9%; Score 2448.8; DB 4; Length 3956;
Best Local Similarity 79.4%; Pred. No. 0;
Matches 3136; Conservative 0; Mismatches 522; Indels 294; Gaps 8;





```
Db       2041 ACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCG 2100
Qy       1811 GTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGC 1870
Db       2101 GTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGC 2160
Qy       1871 CAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGC 1930
Db       2161 CAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGC 2220
Qy       1931 CCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGA 1990
Db       2221 CCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGA 2280
Qy       1991 CTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACC 2050
Db       2281 CTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACC 2340
Qy       2051 CTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCAT 2110
Db       2341 CTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCAT 2400
Qy       2111 AGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTG 2170
Db       2401 AGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTG 2460
Qy       2171 CACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCC 2230
Db       2461 CACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCC 2520
Qy       2231 AACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGA 2290
Db       2521 AACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGA 2580
Qy       2291 GCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACT 2350
Db       2581 GCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACT 2640
Qy       2351 AGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTT 2410
Db       2641 AGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTT 2700
Qy       2411 GGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAG 2470
Db       2701 GGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAG 2760
Qy       2471 CAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGG 2530
Db       2761 CAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGG 2820
Qy       2531 TCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAA 2590
Db       2821 TCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAA 2880
Qy       2591 AGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATA 2650
Db       2881 AGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATA 2940
Qy       2651 TATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCG 2710
Db       2941 TATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCG 3000
Qy       2711 ATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATA 2770
Db       3001 ATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATA 3060
Qy       2771 CGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCG 2830
Db       3061 CGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCG 3120
Qy       2831 GCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCT 2890
Db       3121 GCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCT 3180
Qy       2891 GCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGT 2950
Db       3181 GCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGT 3240
Qy       2951 TCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGC 3010
Db       3241 TCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGC 3300
Qy       3011 TCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGA 3070
Db       3301 TCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGA 3360
Qy       3071 TCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGT 3130
Db       3361 TCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGT 3420
Qy       3131 AAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTC 3190
Db       3421 AAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTC 3480
Qy       3191 ATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAA 3250
Db       3481 ATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAA 3540
Qy       3251 TAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCA 3310
Db       3541 TAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCA 3600
Qy       3311 CATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCA 3370
Db       3601 CATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCA 3660
Qy       3371 AGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCT 3430
Db       3661 AGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCT 3720
Qy       3431 TCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCC 3490
Db       3721 TCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCC 3780
Qy       3491 GCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAA 3550
Db       3781 GCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAA 3840
Qy       3551 TATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATT 3610
Db       3841 TATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATT 3900
Qy       3611 TAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCAC 3662
Db       3901 TAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCAC 3952
```

RESULT 7
US-09-402-266B-20
; Sequence 20, Application US/09402266B
; Patent No. 6537767
; GENERAL INFORMATION:
; APPLICANT: HINNEN, Albert
; APPLICANT: HEGEMANN, Johannes
; APPLICANT: MUNDER, Thomas
; APPLICANT: SCHUSTER, Tilmer
; APPLICANT: FELDMANN, Horst
; APPLICANT: KRAMER, Wilfried
; APPLICANT: ZIMMERMANN, Friedrich
; APPLICANT: ENTIAN, Karl-Dieter
; TITLE OF INVENTION: Process for Screening Antimycotically Active Substances
; FILE REFERENCE: 38005-0094
; CURRENT APPLICATION NUMBER: US/09/402,266B
; CURRENT FILING DATE: 1999-10-08
; PRIOR APPLICATION NUMBER: PCT/EP98/01904

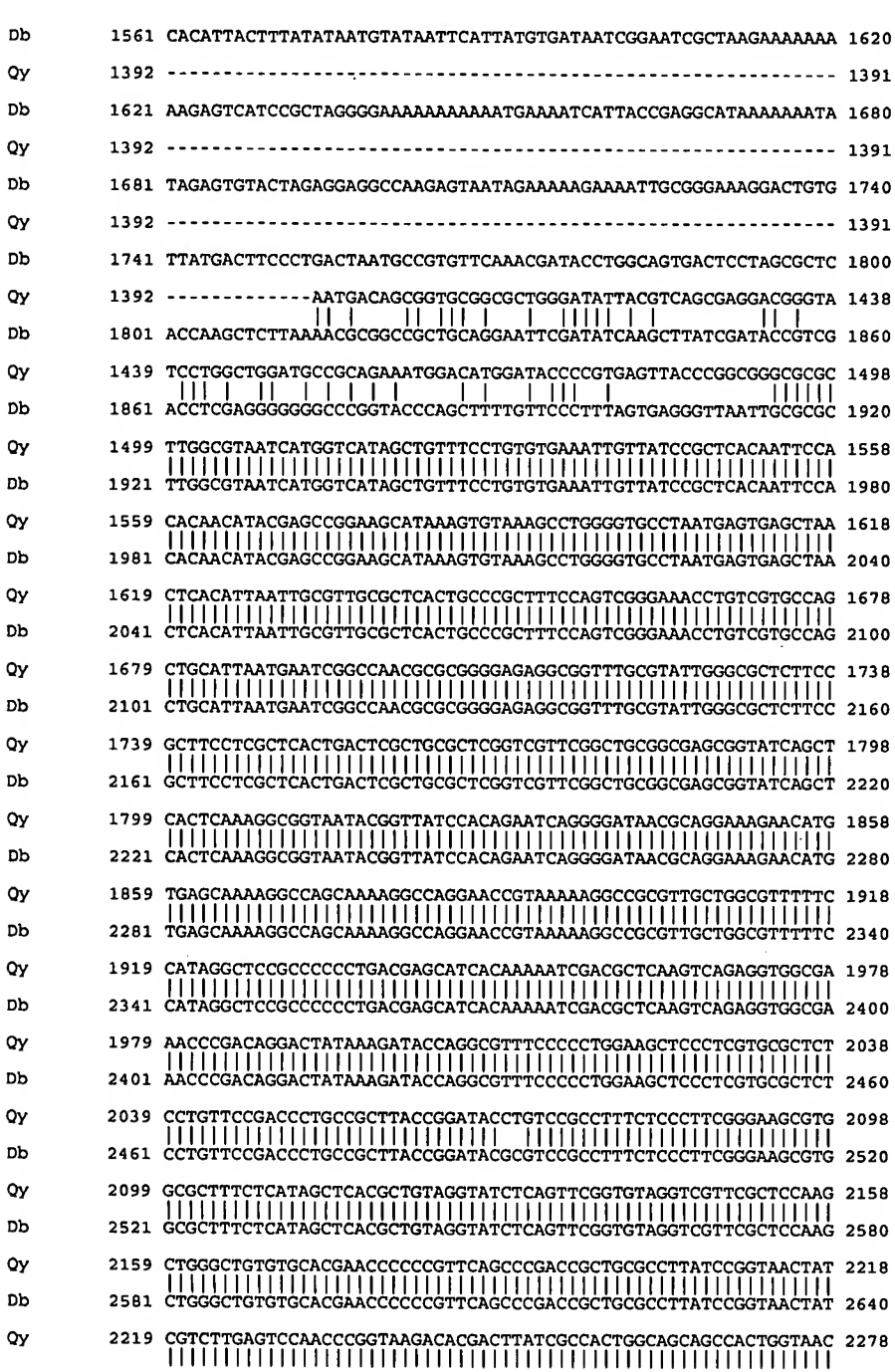
; PRIOR FILING DATE: 1998-10-08
; PRIOR APPLICATION NUMBER: DE 19713572.2
; PRIOR FILING DATE: 1997-04-02
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 20
; LENGTH: 4102
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: ()..()
; OTHER INFORMATION: Synthetic sequence
US-09-402-266B-20

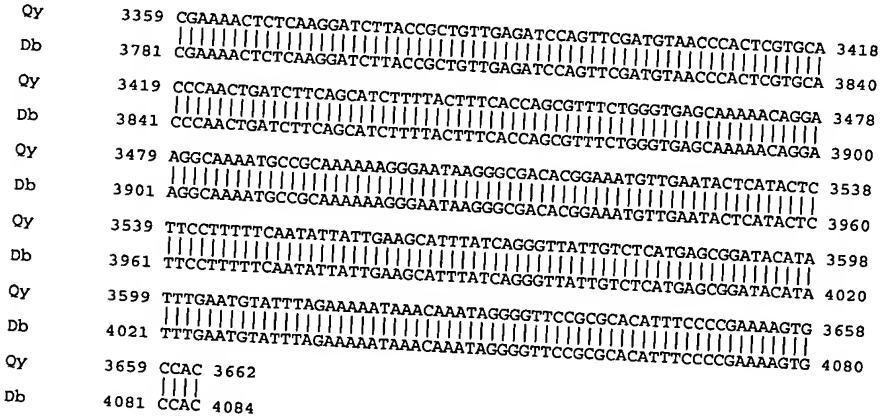
Query Match 62.9%; Score 2305.2; DB 4; Length 4102;
Best Local Similarity 76.9%; Pred. No. 0;
Matches 3150; Conservative 0; Mismatches 508; Indels 440; Gaps 10;

```
Qy        5 ATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTCATTT 64
Db        1 ATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTCATTT 64

Qy       65 TTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGAGATA 124
Db       61 TTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGAGATA 120

Qy      125 GGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTCCAAC 184
Db      121 GGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTCCAAC 180

Qy      185 GTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCCTAA 244
Db      181 GTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCCTAA 240

Qy      245 TCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAGCCCC 304
Db      241 TCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAGCCCC 300

Qy      305 CGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAAAGCG 364
Db      301 CGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAAAGCG 360

Qy      365 AAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCACCACA 424
Db      361 AAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCACCACA 420

Qy      425 CCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCCCATTCGCCATTCAGGCTGCGCAAC 484
Db      421 CCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCCCATTCGCCATTCAGGCTGCGCAAC 480

Qy      485 TGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGGGGA 544
Db      481 TGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGGGGA 540

Qy      545 TGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTGTAAA 604
Db      541 TGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTGTAAA 600

Qy      605 ACGACGGCCAGTGAGCGCGCGTAATACGACTCACTATAGGGCGAATTGGAGCTCCACCGC 660
Db      601 ACGACGGCCAGTGAGCGCGCGTAATACGACTCACTATAGGGCGAATTGGAGCTCCACCGC 664

Qy      665 GGTGGCGCTCTAGAACTAGTGGATCCGGCCATGGAGGCCTTCAAGAATTAGCTTTTCAAT 720
Db      661 GGTGGCGCTCTAGAACTAGTGGATCCGGCCATGGAGGCCTTCAAGAATTAGCTTTTCAAT 720

Qy      721 ---------- 710
Db      721 TCAATTCATCATTTTTTTTTATTCTTTTTTTTGATTTCGGTTTCTTTGAAATTTTTTTG 780

Qy      711 ---------- 710
Db      781 ATTCGGTAATCTCCGAACAGAAGGAAGAACGAAGGAAGGAGCACAGACTTAGATTGGTAT 840

Qy      711 ---------- 710
Db      841 ATATACGCATATGTAGTGTTGAAGAAACATGAAATTGCCCAGTATTCTTAACCCAACTGC 900

Qy      711 --TGCTCGACACGCTGCAGAACACGCAGCTAGATTAACCCTAG---------- 751
Db      901 ACAGAACAAAAACATGCAGGAAACGAAGATAAATCATGTCGAAAGCTACATATAAGGAAC 960

Qy      752 ---------- 756
Db      961 GTGCTGCTACTCATCCTAGTCCTGTTGCTGCCAAGCTATTTAATATCATGCACGAAAAGC 1020

Qy      757 TAATCATATTGTGA----------CGTACGTTAAAGATAATCATGCGTAAAATTG 801
Db     1021 AAACAAACTTGTGTGCTTCATTGGATGTTCGTACCACCAAGGAATTACTGGAGTTAGTTG 1080

Qy      802 ACGCATGGGATCTGTAATACGACTCACTATAGGGCGAATTGGGTACCGGGCCCCCCCTCG 861
Db     1081 AAGCATTAGGTCCCAAAATTTGTTTACTAAAAACACATGTGGATATCTTGACTGATTTTT 1140

Qy      862 AGGTCGACGGTATCGATAAGC---------- 882
Db     1141 CCATGGAGGGCACAGTTAAGCCGCTAAAGGCATTATCCGCCAAGTACAATTTTTTACTCT 1200

Qy      883 ----------TTGATATCGAATTCCTGCAGCCCGGGGGAT 912
Db     1201 TCGAAGACAGAAAATTTGCTGACATTGGTAATACAGTCAAATTGCAGTACTCTGCGGGTG 1260

Qy      913 CCACTAGTTCTAGAGCGGCCGC----------CACCGCGGTGGAGCTCCAGCTT 956
Db     1261 TATACAGAATAGCAGAATGGGCAGACATTACGAATGCACACGGTGTGGTGGGCCCAGGTA 1320

Qy      957 TTGTTCCCTTTAGTGAGGG----------TTAA 979
Db     1321 TTGTTAGCGGTTTGAAGCAGGCGGCAGAAGAAGTAACAAAGGAACCTAGAGGCCTTTTGA 1380

Qy      980 TTAGATCCCATGCGTCAATTTTACGCAGACTATCTTTCTAGGGTTAATCTAGCTGCATCA 1039
Db     1381 TGTTAGCAGAATTGTCATGCAAGGGCTCCCTATCTACTGGAGAATATACTAAGGGTACTG 1440

Qy     1040 GGATCATATCGTCGGG----------TCTTTTTTCCGGCTCAGTCATCGCCCAAGCTGGC 1089
Db     1441 TTGACATTGCGAAGAGCGACAAAGATTTTGTTATCGGCTTTATTGCTCAAAGAGACATGG 1500

Qy     1090 GCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCGGCAT 1149
Db     1501 GTGGAAGAGATGAAGGTTACGATTGGTTGATTATGACACCCGGTGTGGGTTTAGATGACA 1560

Qy     1150 GGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACGTTTA 1209
Db     1561 AGGGAGACGCATTGGGTCAACAGTATAGAACCGTGGATGATGTGGTCTCTACAGGATCTG 1620

Qy     1210 CCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTA----------ACGGTG 1254
Db     1621 ACATTATTATTGTTGGAAGAGGACTATTTGCAAAGGGAAGGGATGCTAAGGTAGAGGGTG 1680

Qy     1255 AACTGTTCGTTCAGGCCACCTGGGATACCAGTT-----CGTCGCGGCTTTTCCGGACACA 1309
Db     1681 AACGTTACAGAAAAGCAGGCTGGGAAGCATATTTGAGAAGATGCGGCCAGCAAAACTAAA 1740

Qy     1310 GTTCCGGATGGTCAGCCCGAAGCGCATCAGCAACCCGAACAATACCGGCGAC-----AGC 1364
Db     1741 AAACTGTATTATAAGTAAATGCATGTATACTAAACTCACAAATTAGAGCTTCAATTTAAT 1800

Qy     1365 CGGAACTGCCGTGCCGGTGTGCAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTC 1424
Db     1801 TATATCAGTTATTACCCAATTCTCATGTTTGCGGCCGCTGCAGGAATTCGATATCAAGCT 1860

Qy     1425 AGCGAGGACGGGTATCCTGGCTGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTA 1484
Db     1861 TATCGATACCGTCGACCTCGAGGGGGGGCCCGGTACCCAGCTTTTGTTCCCTTTAGTGAG 1920
```

```
Qy       1485 CCCGGCGGGCGCGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATC 1544
Db       1921 GGTTAATTGCGCGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATC 1980

Qy       1545 CGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCT 1604
Db       1981 CGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCT 2040

Qy       1605 AATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAA 1664
Db       2041 AATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAA 2100

Qy       1665 ACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTA 1724
Db       2101 ACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTA 2160

Qy       1725 TTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGC 1784
Db       2161 TTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGC 2220

Qy       1785 GAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACG 1844
Db       2221 GAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACG 2280

Qy       1845 CAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGT 1904
Db       2281 CAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGT 2340

Qy       1905 TGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAA 1964
Db       2341 TGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAA 2400

Qy       1965 GTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCT 2024
Db       2401 GTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCT 2460

Qy       2025 CCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCC 2084
Db       2461 CCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCC 2520

Qy       2085 CTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGG 2144
Db       2521 CTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGG 2580

Qy       2145 TCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCT 2204
Db       2581 TCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCT 2640

Qy       2205 TATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAG 2264
Db       2641 TATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAG 2700

Qy       2265 CAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGA 2324
Db       2701 CAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGA 2760

Qy       2325 AGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGA 2384
Db       2761 AGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGA 2820

Qy       2385 AGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTG 2444
Db       2821 AGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTG 2880

Qy       2445 GTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAG 2504
Db       2881 GTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAG 2940

Qy       2505 AAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAG 2564
Db       2941 AAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAG 3000

Qy       2565 GGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAAT 2624
Db       3001 GGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAAT 3060

Qy       2625 GAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCT 2684
Db       3061 GAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCT 3120

Qy       2685 TAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGAC 2744
Db       3121 TAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGAC 3180

Qy       2745 TCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAA 2804
Db       3181 TCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAA 3240

Qy       2805 TGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCG 2864
Db       3241 TGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCG 3300

Qy       2865 GAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATT 2924
Db       3301 GAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATT 3360

Qy       2925 GTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCA 2984
Db       3361 GTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCA 3420

Qy       2985 TTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTT 3044
Db       3421 TTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTT 3480

Qy       3045 CCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCT 3104
Db       3481 CCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCT 3540

Qy       3105 TCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGG 3164
Db       3541 TCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGG 3600

Qy       3165 CAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTG 3224
Db       3601 CAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTG 3660

Qy       3225 AGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGG 3284
Db       3661 AGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGG 3720

Qy       3285 CGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAA 3344
Db       3721 CGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAA 3780

Qy       3345 AACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGT 3404
Db       3781 AACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGT 3840

Qy       3405 AACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGT 3464
Db       3841 AACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGT 3900

Qy       3465 GAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTT 3524
Db       3901 GAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTT 3960

Qy       3525 GAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCA 3584
Db       3961 GAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCA 4020

Qy       3585 TGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACAT 3644
Db       4021 TGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACAT 4080

Qy       3645 TTCCCCGAAAAGTGCCAC 3662
```

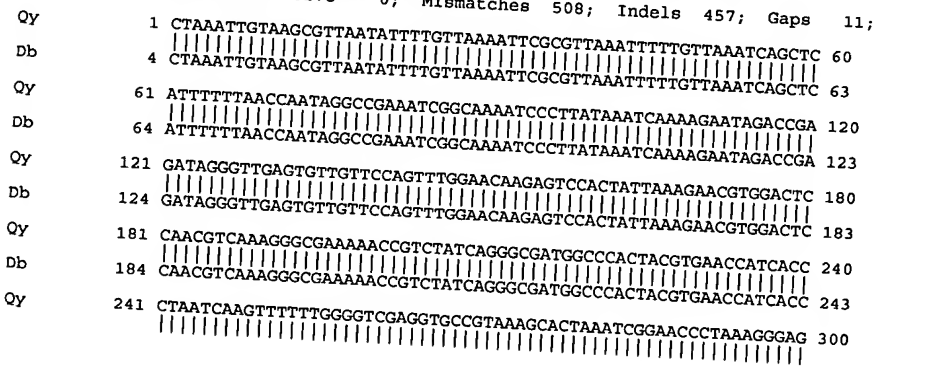
```
Db        4081 TTCCCCGAAAAGTGCCAC 4098

RESULT 8
US-09-402-266B-18
; Sequence 18, Application US/09402266B
; Patent No. 6537767
; GENERAL INFORMATION:
;  APPLICANT: HINNEN, Albert
;  APPLICANT:  HEGEMANN, Johannes
;  APPLICANT:  MUNDER, Thomas
;  APPLICANT:  SCHUSTER, Tilmer
;  APPLICANT:  FELDMANN, Horst
;  APPLICANT:  KRAMER, Wilfried
;  APPLICANT:  ZIMMERMANN, Friedrich
;  APPLICANT:  ENTIAN, Karl-Dieter
;  TITLE OF INVENTION: Process for Screening Antimycotically Active Substances
;  FILE REFERENCE: 38005-0094
;  CURRENT APPLICATION NUMBER: US/09/402,266B
;  CURRENT FILING DATE:  1999-10-08
;  PRIOR APPLICATION NUMBER: PCT/EP98/01904
;  PRIOR FILING DATE: 1998-10-08
;  PRIOR APPLICATION NUMBER: DE 19713572.2
;  PRIOR FILING DATE: 1997-04-02
;  NUMBER OF SEQ ID NOS: 38
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 18
;  LENGTH: 4088
;  TYPE: DNA
;  ORGANISM: Unknown
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: ()..()
;  OTHER INFORMATION: Synthetic sequence
US-09-402-266B-18

  Query Match             62.9%;  Score 2302;  DB 4;  Length 4088;
  Best Local Similarity   76.1%;  Pred. No. 0;
  Matches 3108;  Conservative    0;  Mismatches  550;  Indels  426;  Gaps    5;

Qy          5 ATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTCATTT 64
Db          1 ATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTCATTT 64
Qy         65 TTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGAGATA 124
Db         61 TTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGAGATA 120
Qy        125 GGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTCCAAC 184
Db        121 GGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTCCAAC 180
Qy        185 GTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCCTAA 244
Db        181 GTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCCTAA 240
Qy        245 TCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAGCCCC 304
Db        241 TCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAGCCCC 300
Qy        305 CGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAAAGCG 364
Db        301 CGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAAAGCG 360
Qy        365 AAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCACCACA 424
Db        361 AAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCACCACA 420
Qy        425 CCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCCCATTCGCCATTCAGGCTGCGCAAC 484
Db        421 CCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCCCATTCGCCATTCAGGCTGCGCAAC 480

Qy        485 TGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGGGGA 544
Db        481 TGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGGGGA 540
Qy        545 TGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTGTAAA 604
Db        541 TGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTGTAAA 600
Qy        605 ACGACGGCCAGTGAGCGCGCCTCGTTCATTCACGTTTTTGAACCCGTGGAGGACGGGCAG 664
Db        601 ACGACGGCCAGTGAGCGCGCGTAATACGACTCACTATAGGGCGAATTGGAGCTCCACCGC 660
Qy        665 ACTCGCGGTGCAAATGTGTTTTACAGCGTGATGGAG---------------------- 700
Db        661 GGTGGCGCTCTAGAACTAGTGGATCCGGCCATGGAGGCCTCGTTCAGAATGACACGTATA 720
Qy        701 --CAGATGAAGATGCTCGACACGCTGCAGAACACGCAGCTAGATTAACCCTAGAAAGATA 758
Db        721 GAATGATGCATTACCTTGTCATCTTCAGTATCATACTGTTCGTATACATACTTACTGACA 780
Qy        759 ATCATATTGTGACGTACGTTAAAGATAATCATGCGTAAAATTGACGCATGGGATCTGTAA 818
Db        781 TTCATAGGTATACATATATACACATGTATATATATCGTATGCTGCAGCTTTAAATAATCG 840
Qy        819 TACGACTCACTATAGGGCGAATTGGGTACCGGGCCCCCCCTCGAGGTCGACGGTATCGAT 878
Db        841 GTGTCACTACATAAGAACACCTTTGGTGGAGGGAACATCGTTGGTACCATTGGGCGAGGT 900
Qy        879 AAGCTTGATATCGAATTCCTGCAGCCCGGGGGATCCACTAGTTCTAGAG---------- 927
Db        901 GGCTTCTCTTATGGCAACCGCAAGAGCCTTGAACGCACTCTCACTACGGTGATGATCATT 960
Qy        928 -------CGGCCGCCACCGCGGTGGAGCTCCAGCTTTTGTTCCCTTTAGTGAGGGTTAAT 980
Db        961 CTTGCCTCGCAGACAATCAACGTGGAGGGTAATTCTGCTAGCCTCTGCAAAGCTTTCAAG 1020
Qy        981 TAGATCCCATGCGTCAATTTTACGCAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAG 1040
Db       1021 AAAATGCGGGATCATCTCGCAAGAGAGATCTCCTACTTTCTCCCTTTGCAAACCAAGTTC 1080
Qy       1041 GATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGGC 1100
Db       1081 GACAACTGCGTACGGCCTGTTCGAAAGATCTACCACCGCTCTGGAAAGTGCCTCATCCAA 1140
Qy       1101 ATCGGGGAGGAAGAAGCCCGTGCCTTTT--CCCGCGAGGTTGAAGCGGCATGGAAAGAGT 1158
Db       1141 AGGCGCAAATCCTGATCCAAACCTTTTTACTCCACGCGCCAGTAGGGCCTCTTTAAAAGC 1200
Qy       1159 TTGCCGAGGATGACTGCTGCTGCATTGA---------------------------- 1186
Db       1201 TTGACCGAGAGCAATCCCGCAGTCTTCAGTGGTGTGATGGTCGTCTATGTGTAAGTCACC 1260
Qy       1187 ---------------------------------------------------------- 1186
Db       1261 AATGCACTCAACGATTAGCGACCAGCCGGAATGCTTGGCCAGAGCATGTATCATATGGTC 1320
Qy       1187 CGTTGAGCGAAAACGCACGTTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTT 1246
Db       1321 CAGAAACCCTATACCTGTGTGGACGTTAATCACTTGCGATTGTGTGGCCTGTTCTGCTAC 1380
Qy       1247 TAACGGTGAACTGTTCGTTCAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGAC 1306
Db       1381 TGCTTCTGCCTCTTTTTCTGGGAAGATCGAGTGCTCTATCGCTAGGGGACCACCCTTTAA 1440
Qy       1307 ACAGTTCCGGATGGTCAGCCCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCG 1366
Db       1441 AGAGATCGCAATCTGAATCTTGGTTTCATTTGTAATACGCTTTACTAGGGCTTTCTGCTC 1500
Qy       1367 GAACTGCCGTGCCGGTGTGCAGATT---------------------------------- 1391
Db       1501 TGTCATCTTTGCCTTCGTTTATCTTGCCTGCTCATTTTTTAGTATATTCTTCGAAGAAAT 1560
Qy       1392 ---------------------------------------------------------- 1391
```

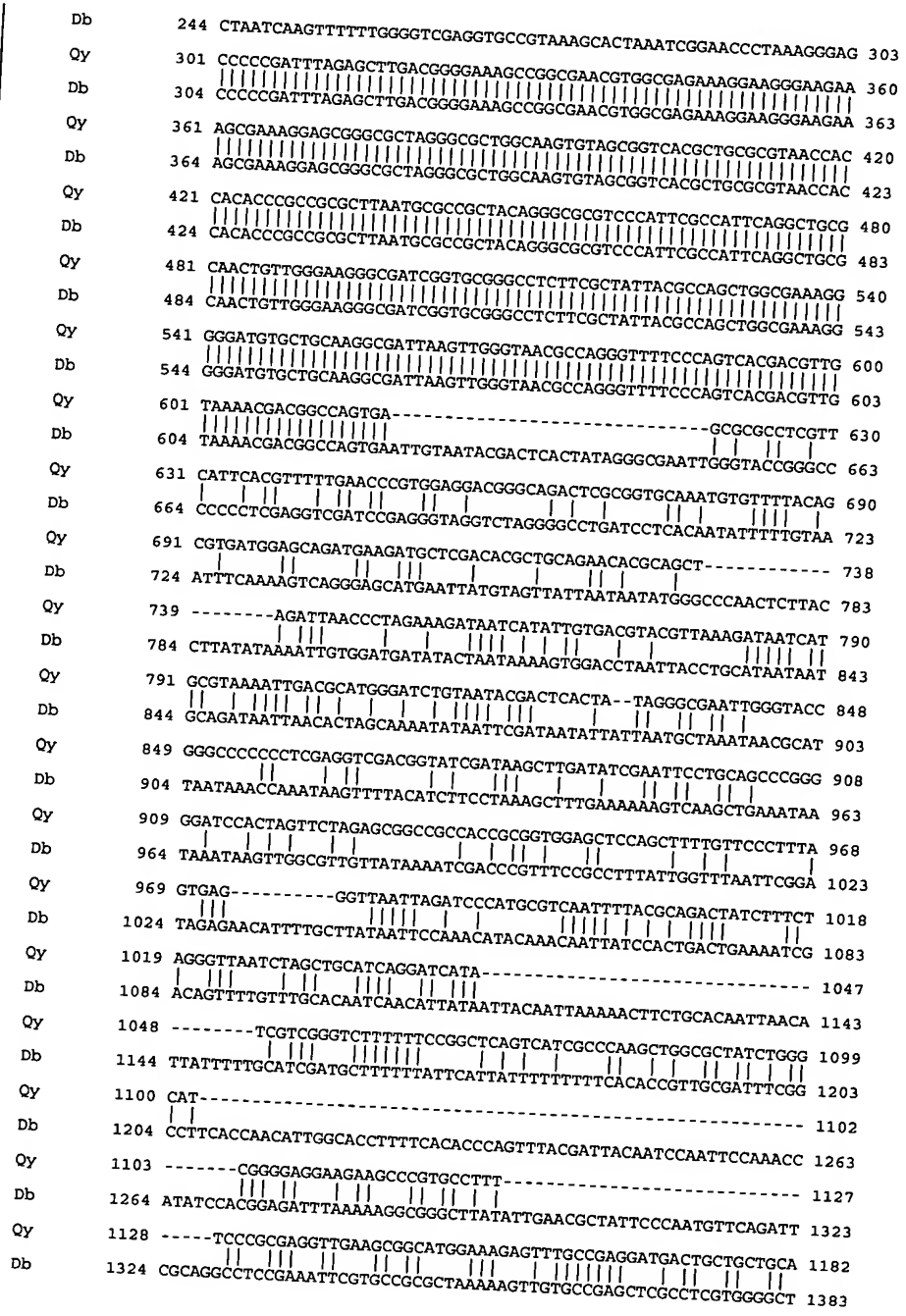

Db 1561 CACATTACTTTATATAATGTATAATTCATTATGTGATAATCGGAATCGCTAAGAAAAAAA 1620
Qy 1392 ---------- 1391
Db 1621 AAGAGTCATCCGCTAGGGGAAAAAAAAAAATGAAAATCATTACCGAGGCATAAAAAAATA 1680
Qy 1392 ---------- 1391
Db 1681 TAGAGTGTACTAGAGGAGGCCAAGAGTAATAGAAAAAGAAAATTGCGGGAAAGGACTGTG 1740
Qy 1392 ---------- 1391
Db 1741 TTATGACTTCCCTGACTAATGCCGTGTTCAAACGATACCTGGCAGTGACTCCTAGCGCTC 1800
Qy 1392 ----------AATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTA 1438
Db 1801 ACCAAGCTCTTAAAACGCGGCCGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCG 1860
Qy 1439 TCCTGGCTGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGC 1498
Db 1861 ACCTCGAGGGGGGGCCCGGTACCCAGCTTTTGTTCCCTTTAGTGAGGGTTAATTGCGCGC 1920
Qy 1499 TTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCA 1558
Db 1921 TTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCA 1980
Qy 1559 CACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAA 1618
Db 1981 CACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAA 2040
Qy 1619 CTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAG 1678
Db 2041 CTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAG 2100
Qy 1679 CTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCC 1738
Db 2101 CTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCC 2160
Qy 1739 GCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCT 1798
Db 2161 GCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCT 2220
Qy 1799 CACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATG 1858
Db 2221 CACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATG 2280
Qy 1859 TGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTC 1918
Db 2281 TGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTC 2340
Qy 1919 CATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGA 1978
Db 2341 CATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGA 2400
Qy 1979 AACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCT 2038
Db 2401 AACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCT 2460
Qy 2039 CCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTG 2098
Db 2461 CCTGTTCCGACCCTGCCGCTTACCGGATACGCGTCCGCCTTTCTCCCTTCGGGAAGCGTG 2520
Qy 2099 GCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAG 2158
Db 2521 GCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAG 2580
Qy 2159 CTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTAT 2218
Db 2581 CTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTAT 2640
Qy 2219 CGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAAC 2278
Db 2641 CGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAAC 2700
Qy 2279 AGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAAC 2338
Db 2701 AGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGACTTCTTGAAGTGGTGGCCTAAC 2760
Qy 2339 TACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTC 2398
Db 2761 TACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTC 2820
Qy 2399 GGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTT 2458
Db 2821 GGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAAGAAACCACCGCTGGTAGCGGTGGTTTT 2880
Qy 2459 TTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATC 2518
Db 2881 TTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATC 2940
Qy 2519 TTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATG 2578
Db 2941 TTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATG 3000
Qy 2579 AGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCA 2638
Db 3001 AGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCA 3060
Qy 2639 ATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCA 2698
Db 3061 ATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCA 3120
Qy 2699 CCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAG 2758
Db 3121 CCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAG 3180
Qy 2759 ATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGAC 2818
Db 3181 ATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGAC 3240
Qy 2819 CCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGC 2878
Db 3241 CCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGC 3300
Qy 2879 AGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCT 2938
Db 3301 AGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCT 3360
Qy 2939 AGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATC 2998
Db 3361 AGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATC 3420
Qy 2999 GTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGG 3058
Db 3421 GTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGG 3480
Qy 3059 CGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATC 3118
Db 3481 CGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATC 3540
Qy 3119 GTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAAT 3178
Db 3541 GTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAAT 3600
Qy 3179 TCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAG 3238
Db 3601 TCTCTTACTGTCATGCCATCCGTAAGATCGTTTTCTGTGACTGGTGAGTACTCAACCAAG 3660
Qy 3239 TCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGAT 3298
Db 3661 TCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGAT 3720
Qy 3299 AATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGG 3358
Db 3721 AATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGG 3780

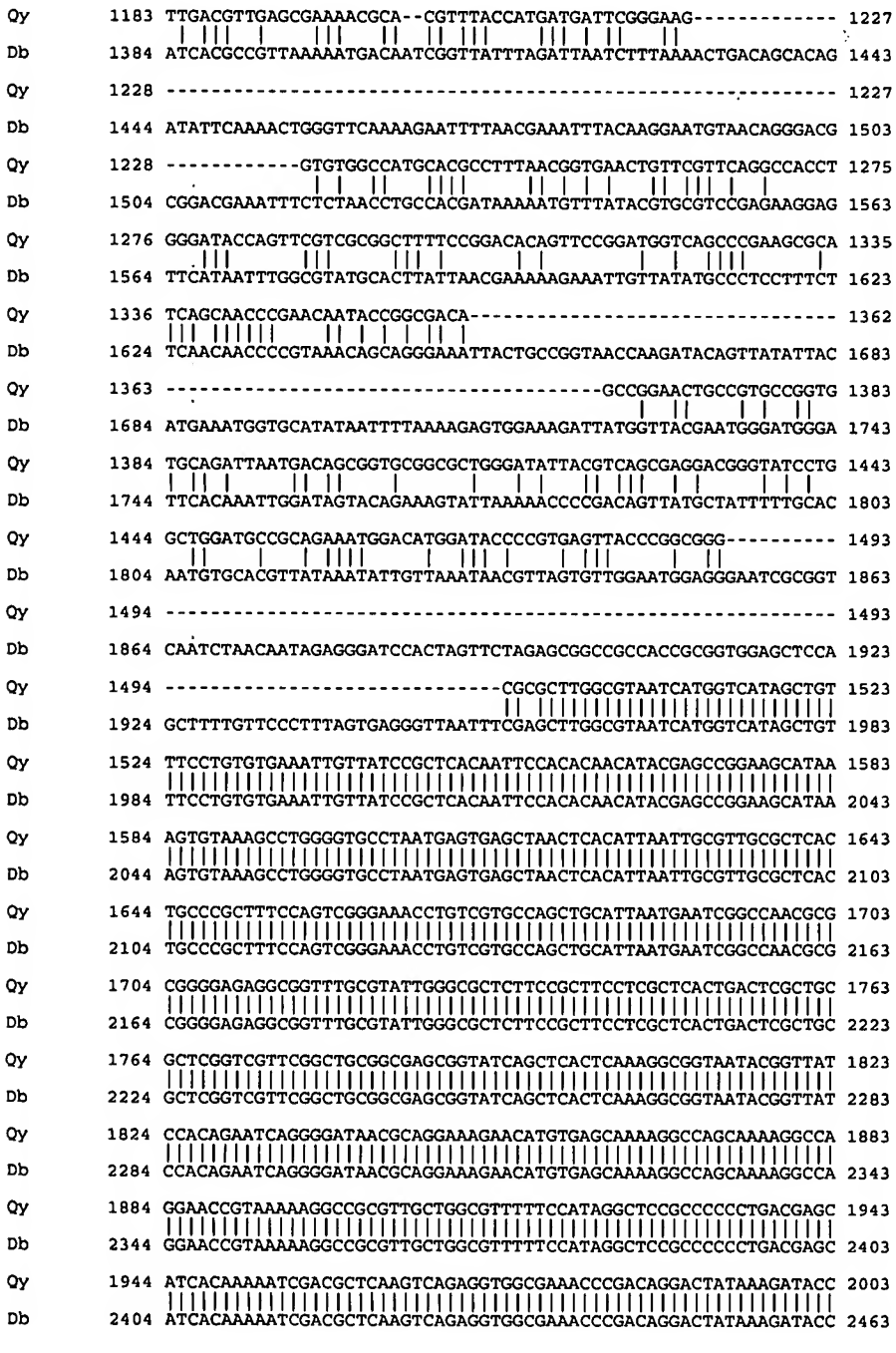


RESULT 9
US-09-993-170-1
; Sequence 1, Application US/09993170
; Patent No. 6743969
; GENERAL INFORMATION:
; APPLICANT: Valent, Barbara
; APPLICANT: Bryan, Gregory
; TITLE OF INVENTION: Modification of a Plant Disease Resistance Gene
; TITLE OF INVENTION: Specificity and Method for Engineering Altered
; TITLE OF INVENTION: Specificity
; FILE REFERENCE: BB1385 US NA
; CURRENT APPLICATION NUMBER: US/09/993,170
; CURRENT FILING DATE: 2001-11-26
; PRIOR APPLICATION NUMBER: 60/248,335
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 1
; LENGTH: 4119
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Plasmid
US-09-993-170-1

Query Match 62.2%; Score 2279.2; DB 4; Length 4119;
Best Local Similarity 76.6%; Pred. No. 0;
Matches 3151; Conservative 0; Mismatches 508; Indels 457; Gaps 11;






Qy 1183 TTGACGTTGAGCGAAAACGCA--CGTTTACCATGATGATTCGGGAAG---------- 1227
Db 1384 ATCACGCCGTTAAAAATGACAATCGGTTATTTAGATTAATCTTTAAAACTGACAGCACAG 1443
Qy 1228 ---------- 1227
Db 1444 ATATTCAAAACTGGGTTCAAAAGAATTTTAACGAAATTTACAAGGAATGTAACAGGGACG 1503
Qy 1228 ----------GTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTTCAGGCCACCT 1275
Db 1504 CGGACGAAATTTCTCTAACCTGCCACGATAAAAATGTTTATACGTGCGTCCGAGAAGGAG 1563
Qy 1276 GGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGCCCGAAGCGCA 1335
Db 1564 TTCATAATTTGGCGTATGCACTTATTAACGAAAAAGAAATTGTTATATGCCCTCCTTTCT 1623
Qy 1336 TCAGCAACCCGAACAATACCGGCGACA---------- 1362
Db 1624 TCAACAACCCCGTAAACAGCAGGGAAATTACTGCCGGTAACCAAGATACAGTTATATTAC 1683
Qy 1363 ----------GCCGGAACTGCCGTGCCGGTG 1383
Db 1684 ATGAAATGGTGCATATAATTTTAAAAGAGTGGAAAGATTATGGTTACGAATGGGATGGGA 1743
Qy 1384 TGCAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTG 1443
Db 1744 TTCACAAATTGGATAGTACAGAAAGTATTAAAAACCCCGACAGTTATGCTATTTTTGCAC 1803
Qy 1444 GCTGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGG---------- 1493
Db 1804 AATGTGCACGTTATAAATATTGTTAAATAACGTTAGTGTTGGAATGGAGGGAATCGCGGT 1863
Qy 1494 ---------- 1493
Db 1864 CAATCTAACAATAGAGGGATCCACTAGTTCTAGAGCGGCCGCCACCGCGGTGGAGCTCCA 1923
Qy 1494 ----------CGCGCTTGGCGTAATCATGGTCATAGCTGT 1523
Db 1924 GCTTTTGTTCCCTTTAGTGAGGGTTAATTTCGAGCTTGGCGTAATCATGGTCATAGCTGT 1983
Qy 1524 TTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAA 1583
Db 1984 TTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAA 2043
Qy 1584 AGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCAC 1643
Db 2044 AGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCAC 2103
Qy 1644 TGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCG 1703
Db 2104 TGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCG 2163
Qy 1704 CGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGC 1763
Db 2164 CGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGC 2223
Qy 1764 GCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTAT 1823
Db 2224 GCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTAT 2283
Qy 1824 CCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCA 1883
Db 2284 CCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCA 2343
Qy 1884 GGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGC 1943
Db 2344 GGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGC 2403
Qy 1944 ATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACC 2003
Db 2404 ATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACC 2463
Qy 2004 AGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCG 2063
Db 2464 AGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCG 2523
Qy 2064 GATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTA 2123
Db 2524 GATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTA 2583
Qy 2124 GGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCG 2183
Db 2584 GGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCG 2643
Qy 2184 TTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGAC 2243
Db 2644 TTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGAC 2703
Qy 2244 ACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAG 2303
Db 2704 ACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAG 2763
Qy 2304 GCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTAT 2363
Db 2764 GCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTAT 2823
Qy 2364 TTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGAT 2423
Db 2824 TTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGAT 2883
Qy 2424 CCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGC 2483
Db 2884 CCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGC 2943
Qy 2484 GCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGT 2543
Db 2944 GCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGT 3003
Qy 2544 GGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCT 2603
Db 3004 GGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCT 3063
Qy 2604 AGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTT 2663
Db 3064 AGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTT 3123
Qy 2664 GGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTC 2723
Db 3124 GGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTC 3183
Qy 2724 GTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTAC 2783
Db 3184 GTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTAC 3243
Qy 2784 CATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTAT 2843
Db 3244 CATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTAT 3303
Qy 2844 CAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCG 2903
Db 3304 CAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCG 3363
Qy 2904 CCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATA 2963
Db 3364 CCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATA 3423
Qy 2964 GTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTA 3023
Db 3424 GTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTA 3483
Qy 3024 TGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGT 3083
Db 3484 TGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGT 3543
Qy 3084 GCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAG 3143

```
Db       3544 GCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAG 3603

Qy       3144 TGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAA 3203

Db       3604 TGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAA 3663

Qy       3204 GATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGC 3263

Db       3664 GATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGC 3723

Qy       3264 GACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTT 3323

Db       3724 GACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTT 3783

Qy       3324 TAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGC 3383

Db       3784 TAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGC 3843

Qy       3384 TGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTA 3443

Db       3844 TGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTA 3903

Qy       3444 CTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAA 3503

Db       3904 CTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAA 3963

Qy       3504 TAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCA 3563

Db       3964 TAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCA 4023

Qy       3564 TTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAAC 3623

Db       4024 TTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAAC 4083

Qy       3624 AAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGC 3659

Db       4084 AAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGC 4119
```

RESULT 10
US-09-138-024-20
; Sequence 20, Application US/09138024A
; Patent No. 6004779
; GENERAL INFORMATION:
; APPLICANT: Bradley, John D.
; APPLICANT: Thompson, Craig M.
; APPLICANT: Moore, Jeffrey B.
; APPLICANT: Wobbe, C. Richard
; APPLICANT: Healy, Judith M.
; APPLICANT: Donnelly, Caroline E.
; TITLE OF INVENTION: REGULATED GENE EXPRESSION IN YEAST
; FILE REFERENCE: 0342/1D469US1
; CURRENT APPLICATION NUMBER: US/09/138,024A
; CURRENT FILING DATE: 1998-08-21
; EARLIER APPLICATION NUMBER: 60/056,719
; EARLIER FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 20
; LENGTH: 7102
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Plasmid pZM195
US-09-138-024-20

Query Match 59.3%; Score 2172.8; DB 3; Length 7102;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2177; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```
Qy       1479 GAGTTACCCGGCGGGCGCGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATT 1538

Db       3341 GAGTCGACCGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATT 3400

Qy       1539 GTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGG 1598

Db       3401 GTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGG 3460

Qy       1599 GTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGT 1658

Db       3461 GTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGT 3520

Qy       1659 CGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTT 1718

Db       3521 CGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTT 3580

Qy       1719 TGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGC 1778

Db       3581 TGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGC 3640

Qy       1779 TGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGG 1838

Db       3641 TGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGG 3700

Qy       1839 ATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGG 1898

Db       3701 ATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGG 3760

Qy       1899 CCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGAC 1958

Db       3761 CCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGAC 3820

Qy       1959 GCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTG 2018

Db       3821 GCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTG 3880

Qy       2019 GAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCT 2078

Db       3881 GAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCT 3940

Qy       2079 TTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGG 2138

Db       3941 TTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGG 4000

Qy       2139 TGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCT 2198

Db       4001 TGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCT 4060

Qy       2199 GCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCAC 2258

Db       4061 GCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCAC 4120

Qy       2259 TGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGT 2318

Db       4121 TGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGT 4180

Qy       2319 TCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTC 2378

Db       4181 TCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTC 4240

Qy       2379 TGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCA 2438

Db       4241 TGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCA 4300

Qy       2439 CCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGAT 2498

Db       4301 CCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGAT 4360

Qy       2499 CTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCAC 2558

Db       4361 CTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCAC 4420

Qy       2559 GTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATT 2618

Db       4421 GTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATT 4480
```

```
Qy       2619 AAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACC 2678
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4481 AAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACC 4540

Qy       2679 AATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTG 2738
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4541 AATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTG 4600

Qy       2739 CCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTG 2798
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4601 CCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTG 4660

Qy       2799 CTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGC 2858
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4661 CTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGC 4720

Qy       2859 CAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTA 2918
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4721 CAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTA 4780

Qy       2919 TTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTG 2978
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4781 TTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTG 4840

Qy       2979 TTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCT 3038
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4841 TTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCT 4900

Qy       3039 CCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTA 3098
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||.
Db       4901 CCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTA 4960

Qy       3099 GCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGG 3158
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4961 GCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGG 5020

Qy       3159 TTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGA 3218
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5021 TTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGA 5080

Qy       3219 CTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTT 3278
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5081 CTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTT 5140

Qy       3279 GCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCA 3338
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5141 GCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCA 5200

Qy       3339 TTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTT 3398
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5201 TTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTT 5260

Qy       3399 CGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTT 3458
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5261 CGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTT 5320

Qy       3459 CTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGA 3518
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5321 CTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGA 5380

Qy       3519 AATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATT 3578
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5381 AATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATT 5440

Qy       3579 GTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGC 3638
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5441 GTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGC 5500

Qy       3639 GCACATTTCCCCGAAAAGTGCCAC 3662
             |||||||||||||||||||||||||
Db       5501 GCACATTTCCCCGAAAAGTGCCAC 5524
```

RESULT 11
US-09-404-066-20
; Sequence 20, Application US/09404066
; Patent No. 6365409
; GENERAL INFORMATION:
; APPLICANT: Bradley, John D.
; APPLICANT: Thompson, Craig M.
; APPLICANT: Moore, Jeffrey B.
; APPLICANT: Wobbe, C. Richard
; APPLICANT: Healy, Judith M.
; APPLICANT: Donnelly, Caroline E.
; TITLE OF INVENTION: REGULATED GENE EXPRESSION IN YEAST
; FILE REFERENCE: 0342/1D469US1
; CURRENT APPLICATION NUMBER: US/09/404,066
; CURRENT FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 09/138,024
; PRIOR FILING DATE: 1998-08-21
; PRIOR APPLICATION NUMBER: 60/056,719
; PRIOR FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 20
; LENGTH: 7102
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Plasmid pZM195
US-09-404-066-20

Query Match 59.3%; Score 2172.8; DB 3; Length 7102;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2177; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```
Qy       1479 GAGTTACCCGGCGGGCGCGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATT 1538
             ||||  |||||  ||  ||||||||||||||||||||||||||||||||||||||||||||
Db       3341 GAGTCGACCGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATT 3400

Qy       1539 GTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGG 1598
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3401 GTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGG 3460

Qy       1599 GTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGT 1658
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3461 GTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGT 3520

Qy       1659 CGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTT 1718
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3521 CGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTT 3580

Qy       1719 TGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGC 1778
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3581 TGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGC 3640

Qy       1779 TGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGG 1838
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3641 TGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGG 3700

Qy       1839 ATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGG 1898
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3701 ATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGG 3760

Qy       1899 CCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGAC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3761 CCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGAC 3820

Qy       1959 GCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3821 GCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTG 3880

Qy       2019 GAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCT 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3881 GAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCT 3940
```

```
Qy       2079 TTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGG 2138
Db       3941 TTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGG 4000
Qy       2139 TGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCT 2198
Db       4001 TGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCT 4060
Qy       2199 GCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCAC 2258
Db       4061 GCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCAC 4120
Qy       2259 TGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGT 2318
Db       4121 TGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGT 4180
Qy       2319 TCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTC 2378
Db       4181 TCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTC 4240
Qy       2379 TGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCA 2438
Db       4241 TGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCA 4300
Qy       2439 CCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGAT 2498
Db       4301 CCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGAT 4360
Qy       2499 CTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCAC 2558
Db       4361 CTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCAC 4420
Qy       2559 GTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATT 2618
Db       4421 GTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATT 4480
Qy       2619 AAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACC 2678
Db       4481 AAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACC 4540
Qy       2679 AATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTG 2738
Db       4541 AATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTG 4600
Qy       2739 CCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTG 2798
Db       4601 CCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTG 4660
Qy       2799 CTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGC 2858
Db       4661 CTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGC 4720
Qy       2859 CAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTA 2918
Db       4721 CAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTA 4780
Qy       2919 TTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTG 2978
Db       4781 TTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTG 4840
Qy       2979 TTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCT 3038
Db       4841 TTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCT 4900
Qy       3039 CCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTA 3098
Db       4901 CCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTA 4960
Qy       3099 GCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGG 3158
Db       4961 GCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGG 5020
Qy       3159 TTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGA 3218
Db       5021 TTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGA 5080
Qy       3219 CTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTT 3278
Db       5081 CTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTT 5140
Qy       3279 GCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCA 3338
Db       5141 GCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCA 5200
Qy       3339 TTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTT 3398
Db       5201 TTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTT 5260
Qy       3399 CGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTT 3458
Db       5261 CGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTT 5320
Qy       3459 CTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGA 3518
Db       5321 CTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGA 5380
Qy       3519 AATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATT 3578
Db       5381 AATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATT 5440
Qy       3579 GTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGC 3638
Db       5441 GTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGC 5500
Qy       3639 GCACATTTCCCCGAAAAGTGCCAC 3662
Db       5501 GCACATTTCCCCGAAAAGTGCCAC 5524
```

RESULT 12
US-09-573-322-20
; Sequence 20, Application US/09573322
; Patent No. 6531289
; GENERAL INFORMATION:
; APPLICANT: Bradley, John D.
; APPLICANT: Thompson, Craig M.
; APPLICANT: Moore, Jeffrey B.
; APPLICANT: Wobbe, C. Richard
; APPLICANT: Bailey, David A.
; TITLE OF INVENTION: Regulated Gene Expression in Yeast and
; TITLE OF INVENTION: Methods of Use
; FILE REFERENCE: 0342/1D469-US4
; CURRENT APPLICATION NUMBER: US/09/573,322
; CURRENT FILING DATE: 2000-05-18
; PRIOR APPLICATION NUMBER: 09/404,066
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 09/138,024
; PRIOR FILING DATE: 1998-08-21
; PRIOR APPLICATION NUMBER: 60/056,719
; PRIOR FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 20
; LENGTH: 7102
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: ZM195 plasmid
US-09-573-322-20

Query Match 59.3%; Score 2172.8; DB 4; Length 7102;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2177; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```
Qy       1479 GAGTTACCCGGCGGGCGCGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATT 1538
```

```
Db      3341 GAGTCGACCGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATT 3400
Qy      1539 GTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGG 1598
Db      3401 GTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGG 3460
Qy      1599 GTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGT 1658
Db      3461 GTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGT 3520
Qy      1659 CGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTT 1718
Db      3521 CGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTT 3580
Qy      1719 TGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGC 1778
Db      3581 TGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGC 3640
Qy      1779 TGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGG 1838
Db      3641 TGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGG 3700
Qy      1839 ATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGG 1898
Db      3701 ATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGG 3760
Qy      1899 CCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGAC 1958
Db      3761 CCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGAC 3820
Qy      1959 GCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTG 2018
Db      3821 GCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTG 3880
Qy      2019 GAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCT 2078
Db      3881 GAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCT 3940
Qy      2079 TTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGG 2138
Db      3941 TTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGG 4000
Qy      2139 TGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCT 2198
Db      4001 TGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCT 4060
Qy      2199 GCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCAC 2258
Db      4061 GCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCAC 4120
Qy      2259 TGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGT 2318
Db      4121 TGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGT 4180
Qy      2319 TCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTC 2378
Db      4181 TCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTC 4240
Qy      2379 TGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCA 2438
Db      4241 TGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCA 4300
Qy      2439 CCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGAT 2498
Db      4301 CCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGAT 4360
Qy      2499 CTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCAC 2558
Db      4361 CTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCAC 4420
Qy      2559 GTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATT 2618
Db      4421 GTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATT 4480
Qy      2619 AAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACC 2678
Db      4481 AAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACC 4540
Qy      2679 AATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTG 2738
Db      4541 AATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTG 4600
Qy      2739 CCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTG 2798
Db      4601 CCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTG 4660
Qy      2799 CTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGC 2858
Db      4661 CTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGC 4720
Qy      2859 CAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTA 2918
Db      4721 CAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTA 4780
Qy      2919 TTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTG 2978
Db      4781 TTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTG 4840
Qy      2979 TTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCT 3038
Db      4841 TTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCT 4900
Qy      3039 CCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTA 3098
Db      4901 CCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTA 4960
Qy      3099 GCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGG 3158
Db      4961 GCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGG 5020
Qy      3159 TTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGA 3218
Db      5021 TTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGA 5080
Qy      3219 CTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTT 3278
Db      5081 CTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTT 5140
Qy      3279 GCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCA 3338
Db      5141 GCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCA 5200
Qy      3339 TTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTT 3398
Db      5201 TTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTT 5260
Qy      3399 CGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTT 3458
Db      5261 CGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTT 5320
Qy      3459 CTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGA 3518
Db      5321 CTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGA 5380
Qy      3519 AATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATT 3578
Db      5381 AATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATT 5440
Qy      3579 GTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGC 3638
Db      5441 GTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGC 5500
Qy      3639 GCACATTTCCCCGAAAAGTGCCAC 3662
Db      5501 GCACATTTCCCCGAAAAGTGCCAC 5524
```

RESULT 13
US-09-138-024-21
; Sequence 21, Application US/09138024A
; Patent No. 6004779
; GENERAL INFORMATION:
; APPLICANT: Bradley, John D.
; APPLICANT: Thompson, Craig M.
; APPLICANT: Moore, Jeffrey B.
; APPLICANT: Wobbe, C. Richard
; APPLICANT: Healy, Judith M.
; APPLICANT: Donnelly, Caroline E.
; TITLE OF INVENTION: REGULATED GENE EXPRESSION IN YEAST
; FILE REFERENCE: 0342/1D469US1
; CURRENT APPLICATION NUMBER: US/09/138,024A
; CURRENT FILING DATE: 1998-08-21
; EARLIER APPLICATION NUMBER: 60/056,719
; EARLIER FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 21
; LENGTH: 7333
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Plasmid pZM197
US-09-138-024-21

Query Match 59.3%; Score 2172.8; DB 3; Length 7333;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2177; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```
Qy       1479 GAGTTACCCGGCGGGCGCGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATT 1538
             |||| |||||  ||  ||||||||||||||||||||||||||||||||||||||||||||
Db       3576 GAGTCGACCGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATT 3635

Qy       1539 GTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGG 1598
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3636 GTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGG 3695

Qy       1599 GTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGT 1658
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3696 GTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGT 3755

Qy       1659 CGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTT 1718
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3756 CGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTT 3815

Qy       1719 TGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGC 1778
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3816 TGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGC 3875

Qy       1779 TGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGG 1838
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3876 TGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGG 3935

Qy       1839 ATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGG 1898
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3936 ATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGG 3995

Qy       1899 CCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGAC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3996 CCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGAC 4055

Qy       1959 GCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4056 GCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTG 4115

Qy       2019 GAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCT 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4116 GAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCT 4175

Qy       2079 TTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGG 2138
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4176 TTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGG 4235

Qy       2139 TGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCT 2198
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4236 TGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCT 4295

Qy       2199 GCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCAC 2258
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4296 GCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCAC 4355

Qy       2259 TGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGT 2318
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4356 TGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGT 4415

Qy       2319 TCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTC 2378
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4416 TCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTC 4475

Qy       2379 TGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCA 2438
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4476 TGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCA 4535

Qy       2439 CCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGAT 2498
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4536 CCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGAT 4595

Qy       2499 CTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCAC 2558
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4596 CTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCAC 4655

Qy       2559 GTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATT 2618
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4656 GTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATT 4715

Qy       2619 AAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACC 2678
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4716 AAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACC 4775

Qy       2679 AATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTG 2738
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4776 AATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTG 4835

Qy       2739 CCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTG 2798
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4836 CCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTG 4895

Qy       2799 CTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGC 2858
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4896 CTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGC 4955

Qy       2859 CAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTA 2918
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4956 CAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTA 5015

Qy       2919 TTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTG 2978
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5016 TTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTG 5075

Qy       2979 TTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCT 3038
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5076 TTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCT 5135

Qy       3039 CCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTA 3098
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5136 CCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTA 5195

Qy       3099 GCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGG 3158
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5196 GCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGG 5255

Qy       3159 TTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGA 3218
```

```
Db       5256 TTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGA 5315

Qy       3219 CTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTT 3278
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5316 CTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTT 5375

Qy       3279 GCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCA 3338
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5376 GCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCA 5435

Qy       3339 TTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTT 3398
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5436 TTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTT 5495

Qy       3399 CGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTT 3458
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5496 CGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTT 5555

Qy       3459 CTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGA 3518
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5556 CTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGA 5615

Qy       3519 AATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATT 3578
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5616 AATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATT 5675

Qy       3579 GTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGC 3638
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5676 GTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGC 5735

Qy       3639 GCACATTTCCCCGAAAAGTGCCAC 3662
             |||||||||||||||||||||||||
Db       5736 GCACATTTCCCCGAAAAGTGCCAC 5759
```

RESULT 14
US-09-404-066-21
; Sequence 21, Application US/09404066
; Patent No. 6365409
; GENERAL INFORMATION:
; APPLICANT: Bradley, John D.
; APPLICANT: Thompson, Craig M.
; APPLICANT: Moore, Jeffrey B.
; APPLICANT: Wobbe, C. Richard
; APPLICANT: Healy, Judith M.
; APPLICANT: Donnelly, Caroline E.
; TITLE OF INVENTION: REGULATED GENE EXPRESSION IN YEAST
; FILE REFERENCE: 0342/1D469US1
; CURRENT APPLICATION NUMBER: US/09/404,066
; CURRENT FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 09/138,024
; PRIOR FILING DATE: 1998-08-21
; PRIOR APPLICATION NUMBER: 60/056,719
; PRIOR FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 21
; LENGTH: 7333
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Plasmid pZM197
US-09-404-066-21

Query Match 59.3%; Score 2172.8; DB 3; Length 7333;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2177; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```
Qy       1479 GAGTTACCCGGCGGGCGCGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATT 1538
             ||||  |||||  ||  ||||||||||||||||||||||||||||||||||||||||||||||
Db       3576 GAGTCGACCGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATT 3635

Qy       1539 GTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGG 1598
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3636 GTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGG 3695

Qy       1599 GTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGT 1658
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3696 GTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGT 3755

Qy       1659 CGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTT 1718
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3756 CGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTT 3815

Qy       1719 TGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGC 1778
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3816 TGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGC 3875

Qy       1779 TGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGG 1838
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3876 TGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGG 3935

Qy       1839 ATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGG 1898
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3936 ATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGG 3995

Qy       1899 CCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGAC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3996 CCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGAC 4055

Qy       1959 GCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4056 GCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTG 4115

Qy       2019 GAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCT 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4116 GAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCT 4175

Qy       2079 TTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGG 2138
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4176 TTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGG 4235

Qy       2139 TGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCT 2198
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4236 TGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCT 4295

Qy       2199 GCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCAC 2258
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4296 GCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCAC 4355

Qy       2259 TGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGT 2318
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4356 TGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGT 4415

Qy       2319 TCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTC 2378
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4416 TCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTC 4475

Qy       2379 TGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCA 2438
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4476 TGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCA 4535

Qy       2439 CCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGAT 2498
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4536 CCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGAT 4595

Qy       2499 CTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCAC 2558
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4596 CTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCAC 4655

Qy       2559 GTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATT 2618
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4656 GTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATT 4715

Qy       2619 AAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACC 2678
```

```
Db       4716 AAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACC 4775
Qy       2679 AATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTG 2738
Db       4776 AATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTG 4835
Qy       2739 CCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTG 2798
Db       4836 CCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTG 4895
Qy       2799 CTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGC 2858
Db       4896 CTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGC 4955
Qy       2859 CAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTA 2918
Db       4956 CAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTA 5015
Qy       2919 TTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTG 2978
Db       5016 TTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTG 5075
Qy       2979 TTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCT 3038
Db       5076 TTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCT 5135
Qy       3039 CCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTA 3098
Db       5136 CCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTA 5195
Qy       3099 GCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGG 3158
Db       5196 GCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGG 5255
Qy       3159 TTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGA 3218
Db       5256 TTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGA 5315
Qy       3219 CTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTT 3278
Db       5316 CTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTT 5375
Qy       3279 GCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCA 3338
Db       5376 GCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCA 5435
Qy       3339 TTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTT 3398
Db       5436 TTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTT 5495
Qy       3399 CGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTT 3458
Db       5496 CGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTT 5555
Qy       3459 CTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGA 3518
Db       5556 CTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGA 5615
Qy       3519 AATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATT 3578
Db       5616 AATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATT 5675
Qy       3579 GTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGC 3638
Db       5676 GTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGC 5735
Qy       3639 GCACATTTCCCCGAAAAGTGCCAC 3662
Db       5736 GCACATTTCCCCGAAAAGTGCCAC 5759

RESULT 15
US-09-573-322-21
; Sequence 21, Application US/09573322
; Patent No. 6531289
; GENERAL INFORMATION:
;  APPLICANT: Bradley, John D.
;  APPLICANT:  Thompson, Craig M.
;  APPLICANT:  Moore, Jeffrey B.
;  APPLICANT:  Wobbe, C. Richard
;  APPLICANT:  Bailey, David A.
;  TITLE OF INVENTION: Regulated Gene Expression in Yeast and
;  TITLE OF INVENTION:  Methods of Use
;  FILE REFERENCE: 0342/1D469-US4
;  CURRENT APPLICATION NUMBER: US/09/573,322
;  CURRENT FILING DATE:  2000-05-18
;  PRIOR APPLICATION NUMBER: 09/404,066
;  PRIOR FILING DATE: 1999-09-23
;  PRIOR APPLICATION NUMBER: 09/138,024
;  PRIOR FILING DATE: 1998-08-21
;  PRIOR APPLICATION NUMBER: 60/056,719
;  PRIOR FILING DATE: 1997-08-22
;  NUMBER OF SEQ ID NOS: 30
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 21
;   LENGTH: 7333
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: ZM197 plasmid
US-09-573-322-21

  Query Match             59.3%;  Score 2172.8;  DB 4;  Length 7333;
  Best Local Similarity   99.7%;  Pred. No. 0;
  Matches 2177;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       1479 GAGTTACCCGGCGGGCGCGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATT 1538
Db       3576 GAGTCGACCGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATT 3635
Qy       1539 GTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGG 1598
Db       3636 GTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGG 3695
Qy       1599 GTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGT 1658
Db       3696 GTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGT 3755
Qy       1659 CGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTT 1718
Db       3756 CGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTT 3815
Qy       1719 TGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGC 1778
Db       3816 TGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGC 3875
Qy       1779 TGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGG 1838
Db       3876 TGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGG 3935
Qy       1839 ATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGG 1898
Db       3936 ATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGG 3995
Qy       1899 CCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGAC 1958
Db       3996 CCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGAC 4055
Qy       1959 GCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTG 2018
Db       4056 GCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTG 4115
Qy       2019 GAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCT 2078
Db       4116 GAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCT 4175
```

```
Qy       2079 TTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGG 2138
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4176 TTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGG 4235

Qy       2139 TGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCT 2198
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4236 TGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCT 4295

Qy       2199 GCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCAC 2258
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4296 GCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCAC 4355

Qy       2259 TGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGT 2318
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4356 TGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGT 4415

Qy       2319 TCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTC 2378
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4416 TCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTC 4475

Qy       2379 TGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCA 2438
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4476 TGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCA 4535

Qy       2439 CCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGAT 2498
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4536 CCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGAT 4595

Qy       2499 CTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCAC 2558
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4596 CTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCAC 4655

Qy       2559 GTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATT 2618
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4656 GTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATT 4715

Qy       2619 AAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACC 2678
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4716 AAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACC 4775

Qy       2679 AATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTG 2738
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4776 AATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTG 4835

Qy       2739 CCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTG 2798
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4836 CCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTG 4895

Qy       2799 CTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGC 2858
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4896 CTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGC 4955

Qy       2859 CAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTA 2918
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4956 CAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTA 5015

Qy       2919 TTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTG 2978
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5016 TTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTG 5075

Qy       2979 TTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCT 3038
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5076 TTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCT 5135

Qy       3039 CCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTA 3098
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5136 CCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTA 5195

Qy       3099 GCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGG 3158
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5196 GCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGG 5255

Qy       3159 TTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGA 3218
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5256 TTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGA 5315

Qy       3219 CTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTT 3278
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5316 CTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTT 5375

Qy       3279 GCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCA 3338
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5376 GCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCA 5435

Qy       3339 TTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTT 3398
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5436 TTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTT 5495

Qy       3399 CGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTT 3458
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5496 CGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTT 5555

Qy       3459 CTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGA 3518
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5556 CTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGA 5615

Qy       3519 AATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATT 3578
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5616 AATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATT 5675

Qy       3579 GTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGC 3638
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5676 GTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGC 5735

Qy       3639 GCACATTTCCCCGAAAAGTGCCAC 3662
             |||||||||||||||||||||||||
Db       5736 GCACATTTCCCCGAAAAGTGCCAC 5759
```

Search completed: January 26, 2005, 12:21:01
Job time : 362.259 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: January 25, 2005, 23:27:35 ; Search time 3011.51 Seconds
(without alignments)
11102.026 Million cell updates/sec

Title: US-10-826-523-40
Perfect score: 707
Sequence: 1 ggatcccatgcgtcaatttt.....aaaattgacgcatgggatcc 707

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Search Notes

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	585.4	82.8	20067	12	XXU02425	U02425 Cloning vec
2	585.4	82.8	20125	12	XXU02426	U02426 Cloning vec
3	585.4	82.8	42529	12	CVU39284	U39284 Cloning vec
4	585.4	82.8	42530	12	CVU39285	U39285 Cloning vec
5	585.4	82.8	42531	12	CVU39286	U39286 Cloning vec
6	585.4	82.8	42704	12	CVU37692	U37692 Cloning vec
7	585.4	82.8	48502	7	LAMCG	J02459 Bacteriopha
8	574.6	81.3	5969	12	AF434923	AF434923 Expressio
9	574.6	81.3	6182	12	AF434925	AF434925 Expressio
10	574.6	81.3	6182	12	AF434926	AF434926 Expressio
11	574.6	81.3	6182	12	AF434927	AF434927 Expressio
12	574.6	81.3	6182	12	AF434931	AF434931 Expressio
13	574.6	81.3	6182	12	AF434932	AF434932 Expressio
14	574.6	81.3	7421	12	AF434924	AF434924 Expressio
15	574.6	81.3	7634	12	AF434928	AF434928 Expressio
16	574.6	81.3	7634	12	AF434929	AF434929 Expressio
17	574.6	81.3	7634	12	AF434930	AF434930 Expressio
18	574.6	81.3	7634	12	AF434933	AF434933 Expressio
19	574.6	81.3	7634	12	AF434934	AF434934 Expressio
20	566.2	80.1	14164	1	AE005330	AE005330 Escherich
c 21	566.2	80.1	26565	6	AX702505	AX702505 Sequence
22	566.2	80.1	46819	6	AR204176	AR204176 Sequence
c 23	566.2	80.1	305325	1	AE016765	AE016765 Escherich
24	564.6	79.9	46897	6	BD184772	BD184772 Nucleic a
25	564.6	79.9	222605	1	AP002555	AP002555 Escherich
c 26	561.4	79.4	22306	6	AR204355	AR204355 Sequence
c 27	561.4	79.4	91772	6	BD184765	BD184765 Nucleic a
c 28	561.4	79.4	286485	1	AP002557	AP002557 Escherich
29	407	57.6	46375	7	AF064539	AF064539 Bacteriop
c 30	395.8	56.0	64137	2	AC101160	AC101160 Mus muscu
c 31	395.8	56.0	179840	2	AC068458	AC068458 Homo sapi
32	395.8	56.0	179871	2	AC024602	AC024602 Homo sapi
33	395.8	56.0	194028	2	AC111077	AC111077 Homo sapi
34	395.8	56.0	206486	2	AC074218	AC074218 Mus muscu
c 35	362.2	51.2	228516	2	AC008594	AC008594 Homo sapi
36	343.2	48.5	68028	2	AC101473	AC101473 Mus muscu
c 37	329.8	46.6	66450	2	AC060821	AC060821 Homo sapi
38	320.4	45.3	63579	2	AC100018	AC100018 Mus muscu
39	320.2	45.3	63690	2	AC125419	AC125419 Homo sapi
c 40	231.8	32.8	29466	1	AE008818	AE008818 Salmonell
41	223	31.5	8134	7	PH2HEADTL	M81255 Bacteriopha
c 42	223	31.5	13264	1	AE005368	AE005368 Escherich
43	223	31.5	13386	1	AE015100	AE015100 Shigella
c 44	223	31.5	34063	6	AR204200	AR204200 Sequence
c 45	223	31.5	44029	6	BD184774	BD184774 Nucleic a

ALIGNMENTS

RESULT 1
XXU02425
LOCUS XXU02425 20067 bp DNA linear SYN 29-JAN-1997
DEFINITION Cloning vector lambda EMBL3, left arm.
ACCESSION U02425
VERSION U02425.1 GI:413791
KEYWORDS .
SOURCE Cloning vector lambda EMBL3
ORGANISM Cloning vector lambda EMBL3
artificial sequences; vectors.
REFERENCE 1 (bases 1 to 20067)
AUTHORS Frischauf,A.M., Lehrach,H., Poustka,A. and Murray,N.
TITLE Lambda replacement vectors carrying polylinker sequences
JOURNAL J. Mol. Biol. 170 (4), 827-842 (1983)
MEDLINE 84064856
PUBMED 6315951
REFERENCE 2 (bases 1 to 20067)
AUTHORS Kitts,P.A.
TITLE CLONTECH Vectors On Disc version 1.3
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 20067)
AUTHORS Kitts,P.A.
TITLE Direct Submission
JOURNAL Submitted (07-OCT-1993) Paul A. Kitts, CLONTECH Laboratories, Inc., 1020 East Meadow Circle, Palo Alto, CA 94303, USA
COMMENT This sequence has been compiled from information in the sequence databases, published literature and other sources. If you suspect there is an error in this sequence, please contact CLONTECH's Technical Service Department at (415) 424-8222 or (800) 662-2566, extension 3 or E-mail TECH@CLONTECH.COM.
FEATURES Location/Qualifiers
source 1. .20067
/organism="Cloning vector lambda EMBL3"
/mol_type="genomic DNA"
/db_xref="taxon:31783"
ORIGIN

Query Match 82.8%; Score 585.4; DB 12; Length 20067;
Best Local Similarity 99.8%; Pred. No. 2.3e-148;
Matches 586; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 45 ATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGC 104

```
Db       3055 ATCCAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGC 3114
Qy        105 TGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCCGTGCCTTTTCCCGCGAGGTTGAAGCG 164
Db       3115 TGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCG 3174
Qy        165 GCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACG 224
Db       3175 GCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACG 3234
Qy        225 TTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTT 284
Db       3235 TTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTT 3294
Qy        285 CAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGC 344
Db       3295 CAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGC 3354
Qy        345 CCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTG 404
Db       3355 CCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTG 3414
Qy        405 CAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGC 464
Db       3415 CAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGC 3474
Qy        465 TGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTC 524
Db       3475 TGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTC 3534
Qy        525 ATTCACGTTTTTGAACCCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGC 584
Db       3535 ATTCACGTTTTTGAACCCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGC 3594
Qy        585 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCAGAACACGCAGCT 631
Db       3595 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCAGAACACGCAGCT 3641
```

```
RESULT 2
XXU02426
LOCUS       XXU02426                20125 bp    DNA     linear   SYN 29-JAN-1997
DEFINITION  Cloning vector lambda EMBL3 SP6/T7, left arm.
ACCESSION   U02426
VERSION     U02426.1  GI:413792
KEYWORDS    .
SOURCE      Cloning vector lambda EMBL3 SP6/T7
  ORGANISM  Cloning vector lambda EMBL3 SP6/T7
            artificial sequences; vectors.
REFERENCE   1  (bases 1 to 20125)
  AUTHORS   Kitts,P.A.
  TITLE     CLONTECH Vectors On Disc version 1.3
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 20125)
  AUTHORS   Kitts,P.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-OCT-1993) Paul A. Kitts, CLONTECH Laboratories, Inc.,
            1020 East Meadow Circle, Palo Alto, CA 94303, USA
COMMENT     This sequence has been compiled from information in the sequence
            databases, published literature and other sources. If you suspect
            there is an error in this sequence, please contact CLONTECH's
            Technical Service Department at (415) 424-8222 or (800) 662-2566,
            extension 3 or E-mail TECH@CLONTECH.COM.
FEATURES             Location/Qualifiers
     source          1. .20125
                     /organism="Cloning vector lambda EMBL3 SP6/T7"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:31784"
ORIGIN

  Query Match             82.8%;  Score 585.4;  DB 12;  Length 20125;
  Best Local Similarity   99.8%;  Pred. No. 2.3e-148;
```

```
  Matches  586;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;
Qy         45 ATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGC 104
Db       3055 ATCCAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGC 3114
Qy        105 TGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCG 164
Db       3115 TGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCG 3174
Qy        165 GCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACG 224
Db       3175 GCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACG 3234
Qy        225 TTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTT 284
Db       3235 TTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTT 3294
Qy        285 CAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGC 344
Db       3295 CAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGC 3354
Qy        345 CCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTG 404
Db       3355 CCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTG 3414
Qy        405 CAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGC 464
Db       3415 CAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGC 3474
Qy        465 TGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTC 524
Db       3475 TGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTC 3534
Qy        525 ATTCACGTTTTTGAACCCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGC 584
Db       3535 ATTCACGTTTTTGAACCCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGC 3594
Qy        585 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCAGAACACGCAGCT 631
Db       3595 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCAGAACACGCAGCT 3641
```

```
RESULT 3
CVU39284
LOCUS       CVU39284                42529 bp    DNA     linear   SYN 13-APR-1996
DEFINITION  Cloning vector TLF97-1, lambda phage lacZ translational fusion
            vector, complete sequence.
ACCESSION   U39284
VERSION     U39284.1  GI:1066304
KEYWORDS    .
SOURCE      Cloning vector TLF97-1
  ORGANISM  Cloning vector TLF97-1
            artificial sequences; vectors.
REFERENCE   1  (bases 1 to 42529)
  AUTHORS   St Pierre,R. and Linn,T.
  TITLE     A refined vector system for the in vitro construction of
            single-copy transcriptional or translational fusions to lacZ
  JOURNAL   Gene 169 (1), 65-68 (1996)
  MEDLINE   96186904
   PUBMED   8635751
REFERENCE   2  (bases 1 to 42529)
  AUTHORS   StPierre,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-OCT-1995) Thomas Linn, Microbiology and Immunology,
            University of Western Ontario, London, Ontario N6A5C1, Canada
FEATURES             Location/Qualifiers
     source          1. .42529
                     /organism="Cloning vector TLF97-1"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:43840"
     terminator      complement(19404. .19489)
                     /note="rrnBT2 terminator"
```

```
     gene            complement(19507. .20760)
                     /gene="lacY"
     CDS             complement(19507. .20760)
                     /gene="lacY"
                     /codon_start=1
                     /transl_table=11
                     /product="lactose permease"
                     /protein_id="AAC53648.1"
                     /db_xref="GI:1066305"
                     /translation="MYYLKNTNFWMFGLFFFFYFFIMGAYFPFFPIWLHDINHISKSD
                     TGIIFAAISLFSLLFQPLFGLLSDKLGLRKYLLWIITGMLVMFAPFFIFIFGPLLQYN
                     ILVGSIVGGIYLGFCFNAGAPAVEAFIEKVSRRSNFEFGRARMFGCVGWALCASIVGI
                     MFTINNQFVFWLGSGCALILAVLLFFAKTDAPSSATVANAVGANHSAFSLKLALELFR
                     QPKLWFLSLYVIGVSCTYDVFDQQFANFFTSFFATGEQGTRVFGYVTTMGELLNASIM
                     FFAPLIINRIGGKNALLLAGTIMSVRIIGSSFATSALEVVILKTLHMFEVPFLLVGCF
                     KYITSQFEVRFSATIYLVCFCFFKQLAMIFMSVLAGNMYESIGFQGAYLVLGLVALGF
                     TLISVFTLSGPGPLSLLRRQVNEVA"
     gene            complement(20812. .23859)
                     /gene="lacZ"
     CDS             complement(20812. .>23859)
                     /gene="lacZ"
                     /note="EcoRI site normally present at the 3' end of lacZ
                     has been removed by an uncharacterized mutation"
                     /codon_start=1
                     /transl_table=11
                     /product="beta-galactosidase"
                     /protein_id="AAC53649.1"
                     /db_xref="GI:1066306"
                     /translation="VVLQRRDWENPGVTQLNRLAAHPPFASWRNSEEARTDRPSQQLR
                     SLNGEWRFAWFPAPEAVPESWLECDLPEADTVVVPSNWQMHGYDAPIYTNVTYPITVN
                     PPFVPTENPTGCYSLTFNVDESWLQEGQTRIIFDGVNSAFHLWCNGRWVGYGQDSRLP
                     SEFDLSAFLRAGENRLAVMVLRWSDGSYLEDQDMWRMSGIFRDVSLLHKPTTQISDFH
                     VATRFNDDFSRAVLEAEVQMCGELRDYLRVTVSLWQGETQVASGTAPFGGEIIDERGG
                     YADRVTLRLNVENPKLWSAEIPNLYRAVVELHTADGTLIEAEACDVGFREVRIENGLL
                     LLNGKPLLIRGVNRHEHHPLHGQVMDEQTMVQDILLMKQNNFNAVRCSHYPNHPLWYT
                     LCDRYGLYVVDEANIETHGMVPMNRLTDDPRWLPAMSERVTRMVQRDRNHPSVIIWSL
                     GNESGHGANHDALYRWIKSVDPSRPVQYEGGGADTTATDIICPMYARVDEDQPFPAVP
                     KWSIKKWLSLPGETRPLILCEYAHAMGNSLGGFAKYWQAFRQYPRLQGGFVWDWVDQS
                     LIKYDENGNPWSAYGGDFGDTPNDRQFCMNGLVFADRTPHPALTEAKHQQQFFQFRLS
                     GQTIEVTSEYLFRHSDNELLHWMVALDGKPLASGEVPLDVAPQGKQLIELPELPQPES
                     AGQLWLTVRVVQPNATAWSEAGHISAWQQWRLAENLSVTLPAASHAIPHLTTSEMDFC
                     IELGNKRWQFNRQSGFLSQMWIGDKKQLLTPLRDQFTRAPLDNDIGVSEATRIDPNAW
                     VERWKAAGHYQAEAALLQCTADTLADAVLITTAHAWQHQGKTLFISRKTYRIDGSGQM
                     AITVDVEVASDTPHPARIGLNCQLAQVAERVNWLGLGPQENYPDRLTAACFDRWDLPL
                     SDMYTPYVFPSENGLRCGTRELNYGPHQWRGDFQFNISRYSQQQLMETSHRHLLHAEE
                     GTWLNIDGFHMGIGGDDSWSPSVSAEFQLSAGRYHYQLVWCQK"
     misc_feature    complement(23874. .23957)
                     /note="multiple cloning site"
     terminator      complement(23973. .24141)
                     /note="rrnBT1 terminator"
     gene            24475. .25665
                     /gene="tetR"
     CDS             24475. .25665
                     /gene="tetR"
                     /codon_start=1
                     /transl_table=11
                     /product="tetracycline resistance protein"
                     /protein_id="AAC53650.1"
                     /db_xref="GI:1066307"
                     /translation="MKSNNALIVILGTVTLDAVGIGLVMPVLPGLLRDIVHSDSIASH
                     YGVLLALYALMQFLCAPVLGALSDRFGRRPVLLASLLGATIDYAIMATTPVLWILYAG
                     RIVAGITGATGAVAGAYIADITDGEDRARHFGLMSACFGVGMVAGPVAGGLLGAISLH
                     APFLAAAVLNGLNLLLGCFLMQESHKGERRPMPLRAFNPVSSFRWARGMTIVAALMTV
                     FFIMQLVGQVPAALWVIFGEDRFRWSATMIGLSLAVFGILHALAQAFVTGPATKRFGE
                     KQAIIAGMAADALGYVLLAFATRGWMAFPIMILLASGGIGMPALQAMLSRQVDDDHQG
                     QLQGSLAALTSLTSITGPLIVTAIYAASASTWNGLAWIVGAALYLVCLPALRRGAWSR
                     ATST"
     misc_feature    33602. .35450
                     /note="imm21 substitution"
ORIGIN

  Query Match           82.8%;  Score 585.4;  DB 12;  Length 42529;
  Best Local Similarity 99.8%;  Pred. No. 2.5e-148;
  Matches  586;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         45 ATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGC 104
               ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3055 ATCCAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGC 3114

Qy        105 TGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCG 164
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3115 TGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCG 3174

Qy        165 GCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACG 224
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3175 GCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACG 3234

Qy        225 TTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTT 284
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3235 TTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTT 3294

Qy        285 CAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGC 344
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3295 CAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGC 3354

Qy        345 CCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTG 404
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3355 CCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTG 3414

Qy        405 CAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGC 464
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3415 CAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGC 3474

Qy        465 TGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTC 524
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3475 TGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTC 3534

Qy        525 ATTCACGTTTTTGAACCCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGC 584
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3535 ATTCACGTTTTTGAACCCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGC 3594

Qy        585 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCAGAACACGCAGCT 631
               |||||||||||||||||||||||||||||||||||||||||||||
Db       3595 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCAGAACACGCAGCT 3641


RESULT 4
CVU39285
LOCUS       CVU39285               42530 bp    DNA     linear   SYN 13-APR-1996
DEFINITION  Cloning vector TLF97-2, phage lambda lacZ translational fusion
            vector, complete sequence.
ACCESSION   U39285
VERSION     U39285.1  GI:1066308
KEYWORDS    .
SOURCE      Cloning vector TLF97-2
  ORGANISM  Cloning vector TLF97-2
            artificial sequences; vectors.
REFERENCE   1  (bases 1 to 42530)
  AUTHORS   St Pierre,R. and Linn,T.
  TITLE     A refined vector system for the in vitro construction of
            single-copy transcriptional or translational fusions to lacZ
  JOURNAL   Gene 169 (1), 65-68 (1996)
  MEDLINE   96186904
   PUBMED   8635751
REFERENCE   2  (bases 1 to 42530)
  AUTHORS   StPierre,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-OCT-1995) Thomas Linn, Microbiology and Immunology,
            University of Western Ontario, London, Ontario N6A5C1, Canada
FEATURES             Location/Qualifiers
     source          1. .42530
                     /organism="Cloning vector TLF97-2"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:43841"
     terminator      complement(19404. .19489)
                     /note="rrnBT2 terminator"
```

```
     gene            complement(19507. .20760)
                     /gene="lacY"
     CDS             complement(19507. .20760)
                     /gene="lacY"
                     /codon_start=1
                     /transl_table=11
                     /product="lactose permease"
                     /protein_id="AAC53651.1"
                     /db_xref="GI:1066309"
                     /translation="MYYLKNTNFWMFGLFFFFYFFIMGAYFPFFPIWLHDINHISKSD
                     TGIIFAAISLFSLLFQPLFGLLSDKLGLRKYLLWIITGMLVMFAPFFIFIFGPLLQYN
                     ILVGSIVGGIYLGFCFNAGAPAVEAFIEKVSRRSNFEFGRARMFGCVGWALCASIVGI
                     MFTINNQFVFWLGSGCALILAVLLFFAKTDAPSSATVANAVGANHSAFSLKLALELFR
                     QPKLWFLSLYVIGVSCTYDVFDQQFANFFTSFFATGEQGTRVFGYVTTMGELLNASIM
                     FFAPLIINRIGGKNALLLAGTIMSVRIIGSSFATSALEVVILKTLHMFEVPFLLVGCF
                     KYITSQFEVRFSATIYLVCFCFFKQLAMIFMSVLAGNMYESIGFQGAYLVLGLVALGF
                     TLISVFTLSGPGPLSLLRRQVNEVA"
     gene            complement(20812. .23859)
                     /gene="lacZ"
     CDS             complement(20812. .>23859)
                     /gene="lacZ"
                     /note="EcoRI site normally present at the 3' end of lacZ
                     has been removed by an uncharacterized mutation"
                     /codon_start=1
                     /transl_table=11
                     /product="beta-galactosidase"
                     /protein_id="AAC53652.1"
                     /db_xref="GI:1066310"
                     /translation="VVLQRRDWENPGVTQLNRLAAHPPFASWRNSEEARTDRPSQQLR
                     SLNGEWRFAWFPAPEAVPESWLECDLPEADTVVVPSNWQMHGYDAPIYTNVTYPITVN
                     PPFVPTENPTGCYSLTFNVDESWLQEGQTRIIFDGVNSAFHLWCNGRWVGYGQDSRLP
                     SEFDLSAFLRAGENRLAVMVLRWSDGSYLEDQDMWRMSGIFRDVSLLHKPTTQISDFH
                     VATRFNDDFSRAVLEAEVQMCGELRDYLRVTVSLWQGETQVASGTAPFGGEIIDERGG
                     YADRVTLRLNVENPKLWSAEIPNLYRAVVELHTADGTLIEAEACDVGFREVRIENGLL
                     LLNGKPLLIRGVNRHEHHPLHGQVMDEQTMVQDILLMKQNNFNAVRCSHYPNHPLWYT
                     LCDRYGLYVVDEANIETHGMVPMNRLTDDPRWLPAMSERVTRMVQRDRNHPSVIIWSL
                     GNESGHGANHDALYRWIKSVDPSRPVQYEGGGADTTATDIICPMYARVDEDQPFPAVP
                     KWSIKKWLSLPGETRPLILCEYAHAMGNSLGGFAKYWQAFRQYPRLQGGFVWDWVDQS
                     LIKYDENGNPWSAYGGDFGDTPNDRQFCMNGLVFADRTPHPALTEAKHQQQFFQFRLS
                     GQTIEVTSEYLFRHSDNELLHWMVALDGKPLASGEVPLDVAPQGKQLIELPELPQPES
                     AGQLWLTVRVVQPNATAWSEAGHISAWQQWRLAENLSVTLPAASHAIPHLTTSEMDFC
                     IELGNKRWQFNRQSGFLSQMWIGDKKQLLTPLRDQFTRAPLDNDIGVSEATRIDPNAW
                     VERWKAAGHYQAEAALLQCTADTLADAVLITTAHAWQHQGKTLFISRKTYRIDGSGQM
                     AITVDVEVASDTPHPARIGLNCQLAQVAERVNWLGLGPQENYPDRLTAACFDRWDLPL
                     SDMYTPYVFPSENGLRCGTRELNYGPHQWRGDFQFNISRYSQQQLMETSHRHLLHAEE
                     GTWLNIDGFHMGIGGDDSWSPSVSAEFQLSAGRYHYQLVWCQK"
     misc_feature    complement(23875. .23958)
                     /note="multiple cloning site"
     terminator      complement(23974. .24142)
                     /note="rrnBT1 terminator"
     gene            24476. .25666
                     /gene="tetR"
     CDS             24476. .25666
                     /gene="tetR"
                     /codon_start=1
                     /transl_table=11
                     /product="tetracycline resistance protein"
                     /protein_id="AAC53653.1"
                     /db_xref="GI:1066311"
                     /translation="MKSNNALIVILGTVTLDAVGIGLVMPVLPGLLRDIVHSDSIASH
                     YGVLLALYALMQFLCAPVLGALSDRFGRRPVLLASLLGATIDYAIMATTPVLWILYAG
                     RIVAGITGATGAVAGAYIADITDGEDRARHFGLMSACFGVGMVAGPVAGGLLGAISLH
                     APFLAAAVLNGLNLLLGCFLMQESHKGERRPMPLRAFNPVSSFRWARGMTIVAALMTV
                     FFIMQLVGQVPAALWVIFGEDRFRWSATMIGLSLAVFGILHALAQAFVTGPATKRFGE
                     KQAIIAGMAADALGYVLLAFATRGWMAFPIMILLASGGIGMPALQAMLSRQVDDDHQG
                     QLQGSLAALTSLTSITGPLIVTAIYAASASTWNGLAWIVGAALYLVCLPALRRGAWSR
                     ATST"
     misc_feature    33603. .35451
                     /note="imm21 substitution"
ORIGIN

  Query Match          82.8%;  Score 585.4;  DB 12;  Length 42530;
  Best Local Similarity  99.8%;  Pred. No. 2.5e-148;
    Matches  586;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         45 ATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGC 104
               ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3055 ATCCAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGC 3114

Qy        105 TGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCG 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3115 TGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCG 3174

Qy        165 GCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACG 224
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3175 GCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACG 3234

Qy        225 TTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTT 284
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3235 TTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTT 3294

Qy        285 CAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGC 344
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3295 CAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGC 3354

Qy        345 CCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTG 404
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3355 CCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTG 3414

Qy        405 CAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGC 464
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3415 CAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGC 3474

Qy        465 TGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTC 524
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3475 TGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTC 3534

Qy        525 ATTCACGTTTTTGAACCCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGC 584
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3535 ATTCACGTTTTTGAACCCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGC 3594

Qy        585 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCAGAACACGCAGCT 631
              |||||||||||||||||||||||||||||||||||||||||||||||
Db       3595 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCAGAACACGCAGCT 3641

RESULT 5
CVU39286
LOCUS       CVU39286                42531 bp    DNA     linear   SYN 13-APR-1996
DEFINITION  Cloning vector TLF97-3, phage lambda lacZ translational fusion
            vector, complete sequence.
ACCESSION   U39286
VERSION     U39286.1  GI:1066312
KEYWORDS    .
SOURCE      Cloning vector TLF97-3
  ORGANISM  Cloning vector TLF97-3
            artificial sequences; vectors.
REFERENCE   1  (bases 1 to 42531)
  AUTHORS   St Pierre,R. and Linn,T.
  TITLE     A refined vector system for the in vitro construction of
            single-copy transcriptional or translational fusions to lacZ
  JOURNAL   Gene 169 (1), 65-68 (1996)
  MEDLINE   96186904
  PUBMED    8635751
REFERENCE   2  (bases 1 to 42531)
  AUTHORS   StPierre,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-OCT-1995) Thomas Linn, Microbiology and Immunology,
            University of Western Ontario, London, Ontario N6A5C1, Canada
FEATURES             Location/Qualifiers
     source          1. .42531
                     /organism="Cloning vector TLF97-3"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:43842"
     terminator      complement(19404. .19489)
                     /note="rrnBT2 terminator"
```

```
     gene            complement(19507. .20760)
                     /gene="lacY"
     CDS             complement(19507. .20760)
                     /gene="lacY"
                     /codon_start=1
                     /transl_table=11
                     /product="lactose permease"
                     /protein_id="AAC53654.1"
                     /db_xref="GI:1066313"
                     /translation="MYYLKNTNFWMFGLFFFFYFFIMGAYFPFFPIWLHDINHISKSD
                     TGIIFAAISLFSLLFQPLFGLLSDKLGLRKYLLWIITGMLVMFAPFFIFIFGPLLQYN
                     ILVGSIVGGIYLGFCFNAGAPAVEAFIEKVSRRSNFEFGRARMFGCVGWALCASIVGI
                     MFTINNQFVFWLGSGCALILAVLLFFAKTDAPSSATVANAVGANHSAFSLKLALELFR
                     QPKLWFLSLYVIGVSCTYDVFDQQFANFFTSFFATGEQGTRVFGYVTTMGELLNASIM
                     FFAPLIINRIGGKNALLLAGTIMSVRIIGSSFATSALEVVILKTLHMFEVPFLLVGCF
                     KYITSQFEVRFSATIYLVCFCFFKQLAMIFMSVLAGNMYESIGFQGAYLVLGLVALGF
                     TLISVFTLSGPGPLSLLRRQVNEVA"
     gene            complement(20812. .23859)
                     /gene="lacZ"
     CDS             complement(20812. .>23859)
                     /gene="lacZ"
                     /note="EcoRI site normally present at the 3' end of lacZ
                     has been removed by an uncharacterized mutation"
                     /codon_start=1
                     /transl_table=11
                     /product="beta-galactosidase"
                     /protein_id="AAC53655.1"
                     /db_xref="GI:1066314"
                     /translation="VVLQRRDWENPGVTQLNRLAAHPPFASWRNSEEARTDRPSQQLR
                     SLNGEWRFAWFPAPEAVPESWLECDLPEADTVVVPSNWQMHGYDAPIYTNVTYPITVN
                     PPFVPTENPTGCYSLTFNVDESWLQEGQTRIIFDGVNSAFHLWCNGRWVGYGQDSRLP
                     SEFDLSAFLRAGENRLAVMVLRWSDGSYLEDQDMWRMSGIFRDVSLLHKPTTQISDFH
                     VATRFNDDFSRAVLEAEVQMCGELRDYLRVTVSLWQGETQVASGTAPFGGEIIDERGG
                     YADRVTLRLNVENPKLWSAEIPNLYRAVVELHTADGTLIEAEACDVGFREVRIENGLL
                     LLNGKPLLIRGVNRHEHHPLHGQVMDEQTMVQDILLMKQNNFNAVRCSHYPNHPLWYT
                     LCDRYGLYVVDEANIETHGMVPMNRLTDDPRWLPAMSERVTRMVQRDRNHPSVIIWSL
                     GNESGHGANHDALYRWIKSVDPSRPVQYEGGGADTTATDIICPMYARVDEDQPFPAVP
                     KWSIKKWLSLPGETRPLILCEYAHAMGNSLGGFAKYWQAFRQYPRLQGGFVWDWVDQS
                     LIKYDENGNPWSAYGGDFGDTPNDRQFCMNGLVFADRTPHPALTEAKHQQQFFQFRLS
                     GQTIEVTSEYLFRHSDNELLHWMVALDGKPLASGEVPLDVAPQGKQLIELPELPQPES
                     AGQLWLTVRVVQPNATAWSEAGHISAWQQWRLAENLSVTLPAASHAIPHLTTSEMDFC
                     IELGNKRWQFNRQSGFLSQMWIGDKKQLLTPLRDQFTRAPLDNDIGVSEATRIDPNAW
                     VERWKAAGHYQAEAALLQCTADTLADAVLITTAHAWQHQGKTLFISRKTYRIDGSGQM
                     AITVDVEVASDTPHPARIGLNCQLAQVAERVNWLGLGPQENYPDRLTAACFDRWDLPL
                     SDMYTPYVFPSENGLRCGTRELNYGPHQWRGDFQFNISRYSQQQLMETSHRHLLHAEE
                     GTWLNIDGFHMGIGGDDSWSPSVSAEFQLSAGRYHYQLVWCQK"
     misc_feature    complement(23876. .23959)
                     /note="multiple cloning site"
     terminator      complement(23975. .24143)
                     /note="rrnBT1 terminator"
     gene            24477. .25667
                     /gene="tetR"
     CDS             24477. .25667
                     /gene="tetR"
                     /codon_start=1
                     /transl_table=11
                     /product="tetracycline resistance protein"
                     /protein_id="AAC53656.1"
                     /db_xref="GI:1066315"
                     /translation="MKSNNALIVILGTVTLDAVGIGLVMPVLPGLLRDIVHSDSIASH
                     YGVLLALYALMQFLCAPVLGALSDRFGRRPVLLASLLGATIDYAIMATTPVLWILYAG
                     RIVAGITGATGAVAGAYIADITDGEDRARHFGLMSACFGVGMVAGPVAGGLLGAISLH
                     APFLAAAVLNGLNLLLGCFLMQESHKGERRPMPLRAFNPVSSFRWARGMTIVAALMTV
                     FFIMQLVGQVPAALWVIFGEDRFRWSATMIGLSLAVFGILHALAQAFVTGPATKRFGE
                     KQAIIAGMAADALGYVLLAFATRGWMAFPIMILLASGGIGMPALQAMLSRQVDDDHQG
                     QLQGSLAALTSLTSITGPLIVTAIYAASASTWNGLAWIVGAALYLVCLPALRRGAWSR
                     ATST"
     misc_feature    33604. .35452
                     /note="imm21 substitution"
ORIGIN

  Query Match           82.8%;  Score 585.4;  DB 12;  Length 42531;
  Best Local Similarity 99.8%;  Pred. No. 2.5e-148;
  Matches  586;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         45 ATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGC 104
               ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3055 ATCCAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGC 3114

Qy        105 TGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCG 164
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3115 TGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCG 3174

Qy        165 GCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACG 224
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3175 GCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACG 3234

Qy        225 TTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTT 284
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3235 TTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTT 3294

Qy        285 CAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGC 344
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3295 CAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGC 3354

Qy        345 CCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTG 404
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3355 CCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTG 3414

Qy        405 CAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGC 464
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3415 CAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGC 3474

Qy        465 TGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTC 524
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3475 TGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTC 3534

Qy        525 ATTCACGTTTTTGAACCCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGC 584
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3535 ATTCACGTTTTTGAACCCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGC 3594

Qy        585 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCAGAACACGCAGCT 631
               |||||||||||||||||||||||||||||||||||||||||||||||
Db       3595 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCAGAACACGCAGCT 3641


RESULT 6
CVU37692
LOCUS       CVU37692               42704 bp    DNA     linear   SYN 13-APR-1996
DEFINITION  Cloning vector lambda TXF97, lacZ transcriptional fusion vector,
            complete sequence.
ACCESSION   U37692
VERSION     U37692.1  GI:1051181
KEYWORDS    .
SOURCE      Cloning vector lambda TXF97
  ORGANISM  Cloning vector lambda TXF97
            artificial sequences; vectors.
REFERENCE   1  (bases 1 to 42704)
  AUTHORS   St Pierre,R. and Linn,T.
  TITLE     A refined vector system for the in vitro construction of
            single-copy transcriptional or translational fusions to lacZ
  JOURNAL   Gene 169 (1), 65-68 (1996)
  MEDLINE   96186904
   PUBMED   8635751
REFERENCE   2  (bases 1 to 42704)
  AUTHORS   StPierre,R. and Linn,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-SEP-1995) Thomas Linn, Microbiology and Immunology,
            University of Western Ontario, London, Ontario N6A5C1, Canada
FEATURES             Location/Qualifiers
     source          1. .42704
                     /organism="Cloning vector lambda TXF97"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:43300"
     misc_feature    1. .42704
                     /note="EcoRI site normally present at this site has been
```

```
misc_feature    removed by an uncharacterized mutation"
                5505
                /note="BamHI site normally present at this site has been
terminator      removed by an uncharacterized mutation"
                complement(19404. .19489)
gene            /note="rrnBT2 terminator"
                complement(19507. .20760)
CDS             /gene="lacY"
                complement(19507. .20760)
                /gene="lacY"
                /codon_start=1
                /transl_table=11
                /product="lactose permease"
                /protein_id="AAC53645.1"
                /db_xref="GI:1051182"
                /translation="MYYLKNTNFWMFGLFFFFYFFIMGAYFPFFPIWLHDINHISKSD
                TGIIFAAISLFSLLFQPLFGLLSDKLGLRKYLLWIITGMLVMFAPFFIFIFGPLLQYN
                ILVGSIVGGIYLGFCFNAGAPAVEAFIEKVSRRSNFEFGRARMFGCVGWALCASIVGI
                MFTINNQFVFWLGSGCALILAVLLFFAKTDAPSSATVANAVGANHSAFSLKLALELFR
                QPKLWFLSLYVIGVSCTYDVFDQQFANFFTSFFATGEQGTRVFGYVTTMGELLNASIM
                FFAPLIINRIGGKNALLLAGTIMSVRIIGSSFATSALEVVILKTLHMFEVPFLLVGCF
                KYITSQFEVRFSATIYLVCFCFFKQLAMIFMSVLAGNMYESIGFQGAYLVLGLVALGF
                TLISVFTLSGPGPLSLLRRQVNEVA"
gene            complement(20812. .23886)
CDS             /gene="lacZ"
                complement(20812. .23886)
                /gene="lacZ"
                /note="EcoRI site normally present at the end of lacZ has
                been removed by an uncharacterized mutation"
                /codon_start=1
                /transl_table=11
                /product="beta-galactosidase"
                /protein_id="AAC53646.1"
                /db_xref="GI:1051183"
                /translation="MTMITDSLAVVLQRRDWENPGVTQLNRLAAHPPFASWRNSEEAR
                TDRPSQQLRSLNGEWRFAWFPAPEAVPESWLECDLPEADTVVVPSNWQMHGYDAPIYT
                NVTYPITVNPPFVPTENPTGCYSLTFNVDESWLQEGQTRIIFDGVNSAFHLWCNGRWV
                GYGQDSRLPSEFDLSAFLRAGENRLAVMVLRWSDGSYLEDQDMWRMSGIFRDVSLLHK
                PTTQISDFHVATRFNDDFSRAVLEAEVQMCGELRDYLRVTVSLWQGETQVASGTAPFG
                GEIIDERGGYADRVTLRLNVENPKLWSAEIPNLYRAVVELHTADGTLIEAEACDVGFR
                EVRIENGLLLLNGKPLLIRGVNRHEHHPLHGQVMDEQTMVQDILLMKQNNFNAVRCSH
                YPNHPLWYTLCDRYGLYVVDEANIETHGMVPMNRLTDDPRWLPAMSERVTRMVQRDRN
                HPSVIIWSLGNESGHGANHDALYRWIKSVDPSRPVQYEGGGADTTATDIICPMYARVD
                EDQPFPAVPKWSIKKWLSLPGETRPLILCEYAHAMGNSLGGFAKYWQAFRQYPRLQGG
                FVWDWVDQSLIKYDENGNPWSAYGGDFGDTPNDRQFCMNGLVFADRTPHPALTEAKHQ
                QQFFQFRLSGQTIEVTSEYLFRHSDNELLHWMVALDGKPLASGEVPLDVAPQGKQLIE
                LPELPQPESAGQLWLTVRVVQPNATAWSEAGHISAWQQWRLAENLSVTLPAASHAIPH
                LTTSEMDFCIELGNKRWQFNRQSGFLSQMWIGDKKQLLTPLRDQFTRAPLDNDIGVSE
                ATRIDPNAWVERWKAAGHYQAEAALLQCTADTLADAVLITTAHAWQHQGKTLFISRKT
                YRIDGSGQMAITVDVEVASDTPHPARIGLNCQLAQVAERVNWLGLGPQENYPDRLTAA
                CFDRWDLPLSDMYTPYVFPSENGLRCGTRELNYGPHQWRGDFQFNISRYSQQQLMETS
                HRHLLHAEEGTWLNIDGFHMGIGGDDSWSPSVSAEFQLSAGRYHYQLVWCQK"
misc_feature    complement(24049. .24132)
terminator      /note="multiple cloning site"
                complement(24148. .24316)
gene            /note="rrnBT1 terminator"
                24650. .25840
CDS             /gene="tetR"
                24650. .25840
                /gene="tetR"
                /note="Description: tetracyline resistance gene;
                tetracycline resistance protein"
                /codon_start=1
                /transl_table=11
                /protein_id="AAC53647.1"
                /db_xref="GI:1051184"
                /translation="MKSNNALIVILGTVTLDAVGIGLVMPVLPGLLRDIVHSDSIASH
                YGVLLALYALMQFLCAPVLGALSDRFGRRPVLLASLLGATIDYAIMATTPVLWILYAG
                RIVAGITGATGAVAGAYIADITDGEDRARHFGLMSACFGVGMVAGPVAGGLLGAISLH
                APFLAAAAVLNGLNLLLGCFLMQESHKGERRPMPLRAFNPVSSFRWARGMTIVAALMTV
                FFIMQLVGQVPAALWVIFGEDRFRWSATMIGLSLAVFGILHALAQAFVTGPATKRFGE
                KQAIIAGMAADALGYVLLAFATRGWMAPPIMILLASGGIGMPALQAMLSRQVDDDHQG
                QLQGSLAALTSLTSITGPLIVTAIYAASASTWNGLAWIVGAALYLVCLPALRRGAWSR
```

```
misc_feature    ATST"
                31747
                /note="EcoRI site normally present at this site has been
misc_feature    removed by an uncharacterized mutation"
                33777. .35625
                /note="imm21 substitution"
misc_feature    39168
                /note="EcoRI site normally present at this site has been
ORIGIN          removed by an uncharacterized mutation"

  Query Match              82.8%;  Score 585.4;  DB 12;  Length 42704;
  Best Local Similarity    99.8%;  Pred. No. 2.5e-148;
  Matches  586;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        45 ATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGC 104
             ||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3055 ATCCAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGC 3114

Qy       105 TGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCG 164
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3115 TGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCG 3174

Qy       165 GCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACG 224
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3175 GCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACG 3234

Qy       225 TTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTT 284
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3235 TTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTT 3294

Qy       285 CAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGC 344
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3295 CAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGC 3354

Qy       345 CCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTG 404
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3355 CCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTG 3414

Qy       405 CAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGC 464
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3415 CAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGC 3474

Qy       465 TGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTC 524
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3475 TGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTC 3534

Qy       525 ATTCACGTTTTTGAACCCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGC 584
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3535 ATTCACGTTTTTGAACCCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGC 3594

Qy       585 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCAGAACACGCAGCT 631
             |||||||||||||||||||||||||||||||||||||||||||||||
Db      3595 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCAGAACACGCAGCT 3641

RESULT 7
LAMCG
LOCUS       LAMCG                    48502 bp    DNA     circular PHG 17-APR-2002
DEFINITION  Bacteriophage lambda, complete genome.
ACCESSION   J02459 M17233 M24325 V00636 X00906
VERSION     J02459.1  GI:215104
KEYWORDS    DNA-binding protein; circular; coat protein; complete genome;
            origin of replication; repressor; unidentified reading frame.
SOURCE      Bacteriophage lambda
  ORGANISM  Bacteriophage lambda
            Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;
            Lambda-like viruses.
REFERENCE   1  (bases 1 to 12)
  AUTHORS   Wu,R. and Taylor,E.
  TITLE     Nucleotide sequence analysis of DNA. II. Complete nucleotide
            sequence of the cohesive ends of bacteriophage lambda DNA
  JOURNAL   J. Mol. Biol. 57 (3), 491-511 (1971)
```

MEDLINE 71209066
PUBMED 4931680
REFERENCE 2 (bases 45493 to 45963)
AUTHORS Imada,M. and Tsugita,A.
TITLE Amino acid sequence of lambda phage endolysin
JOURNAL Nature New Biol. 233, 230-231 (1971)
REFERENCE 3 (sites)
AUTHORS Weigel,P.H., Englund,P.T., Murray,K. and Old,R.W.
TITLE The 3'-terminal nucleotide sequences of bacteriophage lambda DNA
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 70 (4), 1151-1155 (1973)
MEDLINE 73215915
PUBMED 4515613
REFERENCE 4 (bases 38597 to 38672)
AUTHORS Dahlberg,J.E. and Blattner,F.R.
TITLE In vitro transcription products of lambda DNA: Nucleotide sequences and regulatory sites
JOURNAL (in) Fox,C.F. and Robinson,W.S. (Eds.); VIRUS RESEARCH. PROCEEDINGS OF 1973 ICN-UCLA SYMPOSIUM: 533-544; Academic Press, New York (1973)
REFERENCE 5 (bases 37945 to 38027)
AUTHORS Maniatis,T., Ptashne,M., Backman,K., Kield,D., Flashman,S., Jeffrey,A. and Maurer,R.
TITLE Recognition sequences of repressor and polymerase in the operators of bacteriophage lambda
JOURNAL Cell 5 (2), 109-113 (1975)
MEDLINE 75185528
PUBMED 1095210
REFERENCE 6 (bases 35583 to 35600)
AUTHORS Kleid,D.G., Agarwal,K.L. and Khorana,H.G.
TITLE The nucleotide sequence in the promoter region of the gene N in bacteriophage lambda
JOURNAL J. Biol. Chem. 250 (14), 5574-5582 (1975)
MEDLINE 75189495
PUBMED 167018
REFERENCE 7 (bases 35434 to 35618)
AUTHORS Dahlberg,J.E. and Blattner,F.R.
TITLE Sequence of the promoter-operator proximal region of the major leftward RNA of bacteriophage lambda
JOURNAL Nucleic Acids Res. 2 (9), 1441-1458 (1975)
MEDLINE 76031664
PUBMED 1178525
REFERENCE 8 (bases 37945 to 38018)
AUTHORS Maniatis,T., Jeffrey,A. and Kleid,D.G.
TITLE Nucleotide sequence of the rightward operator of phage lambda
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 72 (3), 1184-1188 (1975)
MEDLINE 75158212
PUBMED 1055375
REFERENCE 9 (bases 44588 to 44773)
AUTHORS Sklar,J., Yot,P. and Weissman,S.M.
TITLE Determination of genes, restriction sites, and DNA sequences surrounding the 6S RNA template of bacteriophage lambda
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 72 (5), 1817-1821 (1975)
MEDLINE 75217847
PUBMED 1098044
REFERENCE 10 (bases 37905 to 37989)
AUTHORS Walz,,A., Pirrotta,V. and Ineichen,K.
TITLE Lambda repressor regulates the switch between PR and Prm promoters
JOURNAL Nature 262 (5570), 665-669 (1976)
MEDLINE 76267718
PUBMED 958438
REFERENCE 11 (bases 37946 to 38039)
AUTHORS Smith,G.R., Eisen,H., Reichardt,L. and Hedgepeth,J.
TITLE Deletions of lambda phage locating a prm mutation within the rightward operator
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 73 (3), 712-716 (1976)
MEDLINE 76152323
PUBMED 1062780
REFERENCE 12 (bases 35578 to 35667; 37903 to 38027)
AUTHORS Ptashne,M., Backman,K., Humayun,M.Z., Jeffrey,A., Maurer,R., Meyer,B. and Sauer,R.T.
TITLE Autoregulation and function of a repressor in bacteriophage lambda
JOURNAL Science 194 (4261), 156-161 (1976)
MEDLINE 76271154
PUBMED 959843
REFERENCE 13 (bases 35578 to 35667)
AUTHORS Humayun,Z., Jeffrey,A. and Ptashne,M.
TITLE Completed DNA sequences and organization of repressor-binding sites in the operators of phage lambda
JOURNAL J. Mol. Biol. 112 (2), 265-277 (1977)
MEDLINE 77209970
PUBMED 875019
REFERENCE 14 (bases 38610 to 38732)
AUTHORS Scherer,G., Hobom,G. and Kossel,H.
TITLE DNA base sequence of the po promoter region of phage lamdba
JOURNAL Nature 265 (5590), 117-121 (1977)
MEDLINE 77100320
PUBMED 834253
REFERENCE 15 (bases 38041 to 38241)
AUTHORS Roberts,T.M., Shimatake,H., Brady,C. and Rosenberg,M.
TITLE Sequence of Cro gene of bacteriophage lambda
JOURNAL Nature 270 (5634), 274-275 (1977)
MEDLINE 78071724
PUBMED 593347
REFERENCE 16 (bases 27616 to 28935)
AUTHORS Davies,R.W., Schreier,P.H. and Buchel,D.E.
TITLE Nucleotide sequence of the attachment site of coliphage lambda
JOURNAL Nature 270 (5639), 757-760 (1977)
MEDLINE 78071823
PUBMED 593399
REFERENCE 17 (bases 37206 to 37263; 37914 to 37970)
AUTHORS Humayun,Z.
TITLE DNA sequence at the end of the cI gene in bacteriophage lambda
JOURNAL Nucleic Acids Res. 4 (7), 2137-2143 (1977)
MEDLINE 78011659
PUBMED 909767
REFERENCE 18 (bases 27617 to 27934)
AUTHORS Landy,A. and Ross,W.
TITLE Viral integration and excision: structure of the lambda att sites
JOURNAL Science 197 (4309), 1147-1160 (1977)
MEDLINE 77258934
PUBMED 331474
REFERENCE 19 (bases 39062 to 39170)
AUTHORS Denniston-Thompson,K., Moore,D.D., Kruger,K.E., Furth,M.E. and Blattner,F.R.
TITLE Physical structure of the replication origin of bacteriophage lambda
JOURNAL Science 198 (4321), 1051-1056 (1977)
MEDLINE 78054731
PUBMED 929187
REFERENCE 20 (bases 44467 to 44807)
AUTHORS Sklar,J.L.
TITLE Structure and function of two regions of DNA controlling the synthesis of prokaryotic RNAs
JOURNAL Thesis (1977)
REFERENCE 21 (sites)
AUTHORS Adhya,S. and Gottesman,M.
TITLE Control of transcription termination
JOURNAL Annu. Rev. Biochem. 47, 967-996 (1978)
MEDLINE 78234064
PUBMED 354508
REFERENCE 22 (bases 13 to 72; 48391 to 48502)
AUTHORS Nichols,B.P. and Donelson,J.E.
TITLE 178-Nucleotide sequence surrounding the cos site of bacteriophage lambda DNA
JOURNAL J. Virol. 26 (2), 429-434 (1978)
MEDLINE 78197067
PUBMED 666898
REFERENCE 23 (bases 37938 to 38016; 35589 to 35666)
AUTHORS Flashman,S.M.
TITLE Mutational analysis of the operators of bacteriophage lambda
JOURNAL Mol. Gen. Genet. 166 (1), 61-73 (1978)
MEDLINE 79114073
PUBMED 368570
REFERENCE 24 (bases 37990 to 38982)
AUTHORS Schwarz,E., Scherer,G., Hobom,G. and Kossel,H.
TITLE Nucleotide sequence of cro, cII and part of the O gene in phage

```
            lambda DNA
  JOURNAL   Nature 272 (5652), 410-414 (1978)
  MEDLINE   78135462
   PUBMED   264238
REFERENCE   25 (bases 38212 to 38362)
  AUTHORS   Rosenberg,M., Court,D., Shimatake,H., Brady,C. and Wulff,D.L.
  TITLE     The relationship between function and DNA sequence in an
            intercistronic regulatory region in phage lambda
  JOURNAL   Nature 272 (5652), 414-423 (1978)
  MEDLINE   78135463
   PUBMED   634366
REFERENCE   26 (bases 37224 to 37940)
  AUTHORS   Sauer,R.T.
  TITLE     DNA sequence of the bacteriophage gama cI gene
  JOURNAL   Nature 276 (5685), 301-302 (1978)
  MEDLINE   79053284
   PUBMED   714163
REFERENCE   27 (bases 38597 to 39688)
  AUTHORS   Scherer,G.
  TITLE     Nucleotide sequence of the O gene and of the origin of replication
            in bacteriophage lambda DNA
  JOURNAL   Nucleic Acids Res. 5 (9), 3141-3156 (1978)
  MEDLINE   79033241
   PUBMED   704348
REFERENCE   28 (bases 29711 to 29811; 31043 to 31058)
  AUTHORS   Davies,R.W., Schreier,P.H. and Buchel,D.E.
  TITLE     Determination of the endpoints of partial deletion mutants of the
            attachment site of bacteriophage lambda by DNA sequencing
  JOURNAL   Nucleic Acids Res. 5 (9), 3209-3218 (1978)
  MEDLINE   79033246
   PUBMED   704352
REFERENCE   29 (bases 21661 to 31129)
  AUTHORS   Hoess,R.H. and Landy,A.
  TITLE     Structure of the lambda att sites generated by int-dependent
            deletions
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 75 (11), 5437-5441 (1978)
  MEDLINE   79074833
   PUBMED   364480
REFERENCE   30 (bases 38453 to 38500)

  Query Match             82.8%;  Score 585.4;  DB 7;  Length 48502;
  Best Local Similarity   99.8%;  Pred. No. 2.5e-148;
  Matches  586;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         45 ATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGC 104
               ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3055 ATCCAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGC 3114

Qy        105 TGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCG 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3115 TGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCG 3174

Qy        165 GCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACG 224
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3175 GCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACG 3234

Qy        225 TTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTT 284
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3235 TTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTT 3294

Qy        285 CAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGC 344
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3295 CAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGC 3354

Qy        345 CCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTG 404
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3355 CCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTG 3414

Qy        405 CAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGC 464
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3415 CAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGC 3474

Qy        465 TGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTC 524
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3475 TGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTC 3534

Qy        525 ATTCACGTTTTTGAACCCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGC 584
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3535 ATTCACGTTTTTGAACCCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGC 3594

Qy        585 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCAGAACACGCAGCT 631
              |||||||||||||||||||||||||||||||||||||||||||||||
Db       3595 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCAGAACACGCAGCT 3641


RESULT 8
AF434923
LOCUS       AF434923                5969 bp    DNA     circular SYN 04-APR-2002
DEFINITION  Expression vector pIE1-LUC, complete sequence.
ACCESSION   AF434923
VERSION     AF434923.1  GI:16904152
KEYWORDS    .
SOURCE      Expression vector pIE1-LUC
  ORGANISM  Expression vector pIE1-LUC
            artificial sequences; vectors.
REFERENCE   1  (bases 1 to 5969)
  AUTHORS   Zieler,H. and Huynh,C.Q.
  TITLE     Intron-dependent stimulation of marker gene expression in cultured
            insect cells
  JOURNAL   Insect Mol. Biol. 11 (1), 87-95 (2002)
  MEDLINE   21830665
   PUBMED   11841506
REFERENCE   2  (bases 1 to 5969)
  AUTHORS   Zieler,H. and Huynh,C.Q.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-OCT-2001) Medical Entomology Section, Laboratory of
            Parasitic Diseases, National Institutes of Health, 4 Center Drive
            MSC 0425, Bethesda, MD 20892-0425, USA
FEATURES             Location/Qualifiers
     source          1. .5969
                     /organism="Expression vector pIE1-LUC"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:176965"
                     /note="insect-specific"
     enhancer        97. .614
                     /note="Autographa californica nuclear polyhedrosis virus
                     hr5 enhancer"
     promoter        615. .1201
                     /note="Autographa californica nuclear polyhedrosis virus
                     immediate early promoter"
     CDS             1328. .2980
                     /codon_start=1
                     /product="firefly luciferase"
                     /protein_id="AAL30778.1"
                     /db_xref="GI:16904153"
                     /translation="MEDAKNIKKGPAPFYPLEDGTAGEQLHKAMKRYALVPGTIAFTD
                     AHIEVNITYAEYFEMSVRLAEAMKRYGLNTNHRIVVCSENSLQFFMPVLGALFIGVAV
                     APANDIYNERELLNSMNISQPTVVFVSKKGLQKILNVQKKLPIIQKIIIMDSKTDYQG
                     FQSMYTFVTSHLPPGFNEYDFVPESFDRDKTIALIMNSSGSTGLPKGVALPHRTACVR
                     FSHARDPIFGNQIIPDTAILSVVPFHHGFGMFTTLGYLICGFRVVLMYRFEEELFLRS
                     LQDYKIQSALLVPTLFSFFAKSTLIDKYDLSNLHEIASGGAPLSKEVGEAVAKRFHLP
                     GIRQGYGLTETTSAILITPEGDDKPGAVGKVVPFFEAKVVDLDTGKTLGVNQRGELCV
                     RGPMIMSGYVNNPEATNALIDKDGWLHSGDIAYWDEDEHFFIVDRLKSLIKYKGYQVA
                     PAELESILLQHPNIFDAGVAGLPDDDAGELPAAVVVLEHGKTMTEKEIVDYVASQVTT
                     AKKLRGGVVFVDEVPKGLTGKLDARKIREILIKAKKGGKSKL"
     misc_feature    3146. .3251
                     /note="Heliothis virescens p63 chaperonin gene double
                     poly-adenylation sequence"
     CDS             4168. .5028
                     /codon_start=1
                     /product="beta-lactamase"
                     /protein_id="AAL30779.1"
                     /db_xref="GI:16904154"
                     /translation="MSIQHFRVALIPFFAAFCLPVFAHPETLVKVKDAEDQLGARVGY
                     IELDLNSGKILESFRPEERFPMMSTFKVLLCGAVLSRIDAGQEQLGRRIHYSQNDLVE
                     YSPVTEKHLTDGMTVRELCSAAITMSDNTAANLLLTTIGGPKELTAFLHNMGDHVTRL
```

```
                DRWEPELNEAIPNDERDTTMPVAMATTLRKLLTGELLTLASRQQLIDWMEADKVAGPL
                LRSALPAGWFIADKSGAGERGSRGIIAALGPDGKPSRIVVIYTTGSQATMDERNRQIA
                EIGASLIKHW"
     rep_origin 5316. .5735
                /note="ColE1"
ORIGIN

  Query Match            81.3%;  Score 574.6;  DB 12;  Length 5969;
  Best Local Similarity  99.3%;  Pred. No. 1.7e-145;
  Matches  577;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy         45 ATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGC 104
              ||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3508 ATCCAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGC 3567

Qy        105 TGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCG 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3568 TGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCG 3627

Qy        165 GCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACG 224
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3628 GCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACG 3687

Qy        225 TTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTT 284
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3688 TTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTT 3747

Qy        285 CAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGC 344
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3748 CAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGC 3807

Qy        345 CCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTG 404
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3808 CCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTG 3867

Qy        405 CAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGC 464
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3868 CAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGC 3927

Qy        465 TGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTC 524
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3928 TGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTC 3987

Qy        525 ATTCACGTTTTTGAACCCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGC 584
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3988 ATTCACGTTTTTGAACCCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGC 4047

Qy        585 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCAGAACAC 625
              |||||||||||||||||||||||||||||||||||  | ||
Db       4048 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCATGAAAC 4088


RESULT 9
AF434925
LOCUS       AF434925               6182 bp    DNA     circular SYN 04-APR-2002
DEFINITION  Expression vector 409-FOR, complete sequence.
ACCESSION   AF434925
VERSION     AF434925.1  GI:16904158
KEYWORDS    .
SOURCE      Expression vector 409-FOR
  ORGANISM  Expression vector 409-FOR
            artificial sequences; vectors.
REFERENCE   1  (bases 1 to 6182)
  AUTHORS   Zieler,H. and Huynh,C.Q.
  TITLE     Intron-dependent stimulation of marker gene expression in cultured
            insect cells
  JOURNAL   Insect Mol. Biol. 11 (1), 87-95 (2002)
  MEDLINE   21830665
   PUBMED   11841506
REFERENCE   2  (bases 1 to 6182)
  AUTHORS   Zieler,H. and Huynh,C.Q.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-OCT-2001) Medical Entomology Section, Laboratory of
            Parasitic Diseases, National Institutes of Health, 4 Center Drive
            MSC 0425, Bethesda, MD 20892-0425, USA
FEATURES             Location/Qualifiers
     source          1. .6182
                     /organism="Expression vector 409-FOR"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:176967"
                     /note="insect-specific"
     enhancer        97. .614
                     /note="Autographa californica nuclear polyhedrosis virus
                     hr5 enhancer"
     promoter        615. .1201
                     /note="Autographa californica nuclear polyhedrosis virus
                     immediate early promoter"
     intron          1266. .1398
                     /note="artificial intron"
     CDS             1541. .3193
                     /codon_start=1
                     /product="firefly luciferase"
                     /protein_id="AAL30782.1"
                     /db_xref="GI:16904159"
                     /translation="MEDAKNIKKGPAPFYPLEDGTAGEQLHKAMKRYALVPGTIAFTD
                     AHIEVNITYAEYFEMSVRLAEAMKRYGLNTNHRIVVCSENSLQFFMPVLGALFIGVAV
                     APANDIYNERELLNSMNISQPTVVFVSKKGLQKILNVQKKLPIIQKIIIMDSKTDYQG
                     FQSMYTFVTSHLPPGFNEYDFVPESFDRDKTIALIMNSSGSTGLPKGVALPHRTACVR
                     FSHARDPIFGNQIIPDTAILSVVPFHHGFGMFTTLGYLICGFRVVLMYRFEEELFLRS
                     LQDYKIQSALLVPTLFSFFAKSTLIDKYDLSNLHEIASGGAPLSKEVGEAVAKRFHLP
                     GIRQGYGLTETTSAILITPEGDDKPGAVGKVVPFFEAKVVDLDTGKTLGVNQRGELCV
                     RGPMIMSGYVNNPEATNALIDKDGWLHSGDIAYWDEDEHFFIVDRLKSLIKYKGYQVA
                     PAELESILLQHPNIFDAGVAGLPDDDAGELPAAVVVLEHGKTMTEKEIVDYVASQVTT
                     AKKLRGGVVFVDEVPKGLTGKLDARKIREILIKAKKGGKSKL"
     misc_feature    3359. .3464
                     /note="Heliothis virescens p63 chaperonin gene double
                     poly-adenylation sequence"
     CDS             4381. .5241
                     /codon_start=1
                     /product="beta-lactamase"
                     /protein_id="AAL30783.1"
                     /db_xref="GI:16904160"
                     /translation="MSIQHFRVALIPFFAAFCLPVFAHPETLVKVKDAEDQLGARVGY
                     IELDLNSGKILESFRPEERFPMMSTFKVLLCGAVLSRIDAGQEQLGRRIHYSQNDLVE
                     YSPVTEKHLTDGMTVRELCSAAITMSDNTAANLLLTTIGGPKELTAFLHNMGDHVTRL
                     DRWEPELNEAIPNDERDTTMPVAMATTLRKLLTGELLTLASRQQLIDWMEADKVAGPL
                     LRSALPAGWFIADKSGAGERGSRGIIAALGPDGKPSRIVVIYTTGSQATMDERNRQIA
                     EIGASLIKHW"
     rep_origin      5529. .5948
                     /note="ColE1"
ORIGIN

  Query Match            81.3%;  Score 574.6;  DB 12;  Length 6182;
  Best Local Similarity  99.3%;  Pred. No. 1.7e-145;
  Matches  577;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy         45 ATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGC 104
              ||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3721 ATCCAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGC 3780

Qy        105 TGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCG 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3781 TGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCG 3840

Qy        165 GCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACG 224
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3841 GCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACG 3900

Qy        225 TTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTT 284
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3901 TTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTT 3960

Qy        285 CAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGC 344
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3961 CAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGC 4020
```

```
Qy        345 CCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTG 404
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4021 CCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTG 4080

Qy        405 CAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGC 464
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4081 CAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGC 4140

Qy        465 TGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTC 524
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4141 TGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTC 4200

Qy        525 ATTCACGTTTTTGAACCCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGC 584
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4201 ATTCACGTTTTTGAACCCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGC 4260

Qy        585 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCAGAACAC 625
              |||||||||||||||||||||||||||||||||||||  | ||
Db       4261 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCATGAAAC 4301
```

```
RESULT 10
AF434926
LOCUS       AF434926               6182 bp    DNA     circular SYN 04-APR-2002
DEFINITION  Expression vector 409-MUT, complete sequence.
ACCESSION   AF434926
VERSION     AF434926.1  GI:16904161
KEYWORDS    .
SOURCE      Expression vector 409-MUT
  ORGANISM  Expression vector 409-MUT
            artificial sequences; vectors.
REFERENCE   1  (bases 1 to 6182)
  AUTHORS   Zieler,H. and Huynh,C.Q.
  TITLE     Intron-dependent stimulation of marker gene expression in cultured
            insect cells
  JOURNAL   Insect Mol. Biol. 11 (1), 87-95 (2002)
  MEDLINE   21830665
   PUBMED   11841506
REFERENCE   2  (bases 1 to 6182)
  AUTHORS   Zieler,H. and Huynh,C.Q.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-OCT-2001) Medical Entomology Section, Laboratory of
            Parasitic Diseases, National Institutes of Health, 4 Center Drive
            MSC 0425, Bethesda, MD 20892-0425, USA
FEATURES             Location/Qualifiers
     source          1. .6182
                     /organism="Expression vector 409-MUT"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:176968"
     enhancer        /note="insect-specific"
                     97. .614
                     /note="Autographa californica nuclear polyhedrosis virus
                     hr5 enhancer"
     promoter        615. .1201
                     /note="Autographa californica nuclear polyhedrosis virus
                     immediate early promoter"
     intron          1266. .1398
                     /note="artificial intron with mutated 3' splice site"
     CDS             1541. .3193
                     /codon_start=1
                     /product="firefly luciferase"
                     /protein_id="AAL30784.1"
                     /db_xref="GI:16904162"
                     /translation="MEDAKNIKKGPAPFYPLEDGTAGEQLHKAMKRYALVPGTIAFTD
                     AHIEVNITYAEYFEMSVRLAEAMKRYGLNTNHRIVVCSENSLQFFMPVLGALFIGVAV
                     APANDIYNERELLNSMNISQPTVVFVSKKGLQKILNVQKKLPIIQKIIIMDSKTDYQG
                     FQSMYTFVTSHLPPGFNEYDFVPESFDRDKTIALIMNSSGSTGLPKGVALPHRTACVR
                     FSHARDPIFGNQIIPDTAILSVVPFHHGFGMFTTLGYLICGFRVVLMYRFEEELFLRS
                     LQDYKIQSALLVPTLFSFFAKSTLIDKYDLSNLHEIASGGAPLSKEVGEAVAKRFHLP
                     GIRQGYGLTETTSAILITPEGDDKPGAVGKVVPFFEAKVVDLDTGKTLGVNQRGELCV
                     RGPMIMSGYVNNPEATNALIDKDGWLHSGDIAYWDEDEHFFIVDRLKSLIKYKGYQVA
                     PAELESILLQHPNIFDAGVAGLPDDDAGELPAAVVVLEHGKTMTEKEIVDYVASQVTT
                     AKKLRGGVVFVDEVPKGLTGKLDARKIREILIKAKKGGKSKL"
```

```
     misc_feature    3359. .3464
                     /note="Heliothis virescens p63 chaperonin gene double
                     poly-adenylation sequence"
     CDS             4381. .5241
                     /codon_start=1
                     /product="beta-lactamase"
                     /protein_id="AAL30785.1"
                     /db_xref="GI:16904163"
                     /translation="MSIQHFRVALIPFFAAFCLPVFAHPETLVKVKDAEDQLGARVGY
                     IELDLNSGKILESFRPEERFPMMSTFKVLLCGAVLSRIDAGQEQLGRRIHYSQNDLVE
                     YSPVTEKHLTDGMTVRELCSAAITMSDNTAANLLLTTIGGPKELTAFLHNMGDHVTRL
                     DRWEPELNEAIPNDERDTTMPVAMATTLRKLLTGELLTLASRQQLIDWMEADKVAGPL
                     LRSALPAGWFIADKSGAGERGSRGIIAALGPDGKPSRIVVIYTTGSQATMDERNRQIA
                     EIGASLIKHW"
     rep_origin      5529. .5948
                     /note="ColE1"
ORIGIN
```

```
  Query Match             81.3%;  Score 574.6;  DB 12;  Length 6182;
  Best Local Similarity   99.3%;  Pred. No. 1.7e-145;
  Matches  577;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy         45 ATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGC 104
              ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3721 ATCCAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGC 3780

Qy        105 TGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCG 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3781 TGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCG 3840

Qy        165 GCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACG 224
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3841 GCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACG 3900

Qy        225 TTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTT 284
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3901 TTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTT 3960

Qy        285 CAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGC 344
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3961 CAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGC 4020

Qy        345 CCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTG 404
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4021 CCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTG 4080

Qy        405 CAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGC 464
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4081 CAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGC 4140

Qy        465 TGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTC 524
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4141 TGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTC 4200

Qy        525 ATTCACGTTTTTGAACCCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGC 584
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4201 ATTCACGTTTTTGAACCCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGC 4260

Qy        585 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCAGAACAC 625
              |||||||||||||||||||||||||||||||||||||  | ||
Db       4261 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCATGAAAC 4301
```

```
RESULT 11
AF434927
LOCUS       AF434927               6182 bp    DNA     circular SYN 04-APR-2002
DEFINITION  Expression vector 410-FOR, complete sequence.
ACCESSION   AF434927
VERSION     AF434927.1  GI:16904164
KEYWORDS    .
SOURCE      Expression vector 410-FOR
  ORGANISM  Expression vector 410-FOR
            artificial sequences; vectors.
```

```
REFERENCE   1  (bases 1 to 6182)
  AUTHORS   Zieler,H. and Huynh,C.Q.
  TITLE     Intron-dependent stimulation of marker gene expression in cultured
            insect cells
  JOURNAL   Insect Mol. Biol. 11 (1), 87-95 (2002)
  MEDLINE   21830665
   PUBMED   11841506
REFERENCE   2  (bases 1 to 6182)
  AUTHORS   Zieler,H. and Huynh,C.Q.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-OCT-2001) Medical Entomology Section, Laboratory of
            Parasitic Diseases, National Institutes of Health, 4 Center Drive
            MSC 0425, Bethesda, MD 20892-0425, USA
FEATURES             Location/Qualifiers
     source          1. .6182
                     /organism="Expression vector 410-FOR"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:176969"
                     /note="insect-specific"
     enhancer        97. .614
                     /note="Autographa californica nuclear polyhedrosis virus
                     hr5 enhancer"
     promoter        615. .1201
                     /note="Autographa californica nuclear polyhedrosis virus
                     immediate early promoter"
     CDS             1328. .2980
                     /codon_start=1
                     /product="firefly luciferase"
                     /protein_id="AAL30786.1"
                     /db_xref="GI:16904165"
                     /translation="MEDAKNIKKGPAPFYPLEDGTAGEQLHKAMKRYALVPGTIAFTD
                     AHIEVNITYAEYFEMSVRLAEAMKRYGLNTNHRIVVCSENSLQFFMPVLGALFIGVAV
                     APANDIYNERELLNSMNISQPTVVFVSKKGLQKILNVQKKLPIIQKIIIMDSKTDYQG
                     FQSMYTFVTSHLPPGFNEYDFVPESFDRDKTIALIMNSSGSTGLPKGVALPHRTACVR
                     FSHARDPIFGNQIIPDTAILSVVPFHHGFGMFTTLGYLICGFRVVLMYRFEEELFLRS
                     LQDYKIQSALLVPTLFSFFAKSTLIDKYDLSNLHEIASGGAPLSKEVGEAVAKRFHLP
                     GIRQGYGLTETTSAILITPEGDDKPGAVGKVVPFFEAKVVDLDTGKTLGVNQRGELCV
                     RGPMIMSGYVNNPEATNALIDKDGWLHSGDIAYWDEDEHFFIVDRLKSLIKYKGYQVA
                     PAELESILLQHPNIFDAGVAGLPDDDAGELPAAVVVLEHGKTMTEKEIVDYVASQVTT
                     AKKLRGGVVFVDEVPKGLTGKLDARKIREILIKAKKGGKSKL"
     intron          3196. .3328
                     /note="artificial intron"
     misc_feature    3359. .3464
                     /note="Heliothis virescens p63 chaperonin gene double
                     poly-adenylation sequence"
     CDS             4381. .5241
                     /codon_start=1
                     /product="beta-lactamase"
                     /protein_id="AAL30787.1"
                     /db_xref="GI:16904166"
                     /translation="MSIQHFRVALIPFFAAFCLPVFAHPETLVKVKDAEDQLGARVGY
                     IELDLNSGKILESFRPEERFPMMSTFKVLLCGAVLSRIDAGQEQLGRRIHYSQNDLVE
                     YSPVTEKHLTDGMTVRELCSAAITMSDNTAANLLLTTIGGPKELTAFLHNMGDHVTRL
                     DRWEPELNEAIPNDERDTTMPVAMATTLRKLLTGELLTLASRQQLIDWMEADKVAGPL
                     LRSALPAGWFIADKSGAGERGSRGIIAALGPDGKPSRIVVIYTTGSQATMDERNRQIA
                     EIGASLIKHW"
     rep_origin      5529. .5948
                     /note="ColE1"
ORIGIN

  Query Match             81.3%;  Score 574.6;  DB 12;  Length 6182;
  Best Local Similarity   99.3%;  Pred. No. 1.7e-145;
  Matches  577;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy         45 ATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGC 104
              ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3721 ATCCAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGC 3780

Qy        105 TGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCG 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3781 TGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCG 3840

Qy        165 GCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACG 224
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3841 GCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACG 3900

Qy        225 TTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTT 284
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3901 TTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTT 3960

Qy        285 CAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGC 344
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3961 CAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGC 4020

Qy        345 CCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTG 404
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4021 CCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTG 4080

Qy        405 CAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGC 464
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4081 CAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGC 4140

Qy        465 TGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTC 524
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4141 TGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTC 4200

Qy        525 ATTCACGTTTTTGAACCCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGC 584
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4201 ATTCACGTTTTTGAACCCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGC 4260

Qy        585 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCAGAACAC 625
              |||||||||||||||||||||||||||||||||||  | ||
Db       4261 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCATGAAAC 4301


RESULT 12
AF434931
LOCUS       AF434931               6182 bp    DNA     circular SYN 04-APR-2002
DEFINITION  Expression vector 409-REV, complete sequence.
ACCESSION   AF434931
VERSION     AF434931.1  GI:16904176
KEYWORDS    .
SOURCE      Expression vector 409-REV
  ORGANISM  Expression vector 409-REV
            artificial sequences; vectors.
REFERENCE   1  (bases 1 to 6182)
  AUTHORS   Zieler,H. and Huynh,C.Q.
  TITLE     Intron-dependent stimulation of marker gene expression in cultured
            insect cells
  JOURNAL   Insect Mol. Biol. 11 (1), 87-95 (2002)
  MEDLINE   21830665
   PUBMED   11841506
REFERENCE   2  (bases 1 to 6182)
  AUTHORS   Zieler,H. and Huynh,C.Q.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-OCT-2001) Medical Entomology Section, Laboratory of
            Parasitic Diseases, National Institutes of Health, 4 Center Drive
            MSC 0425, Bethesda, MD 20892-0425, USA
FEATURES             Location/Qualifiers
     source          1. .6182
                     /organism="Expression vector 409-REV"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:176973"
                     /note="insect-specific"
     enhancer        97. .614
                     /note="Autographa californica nuclear polyhedrosis virus
                     hr5 enhancer"
     promoter        615. .1201
                     /note="Autographa californica nuclear polyhedrosis virus
                     immediate early promoter"
     intron          complement(1236. .1368)
                     /note="artificial intron"
     CDS             1541. .3193
                     /codon_start=1
                     /product="firefly luciferase"
                     /protein_id="AAL30794.1"
```

```
                     /db_xref="GI:16904177"
                     /translation="MEDAKNIKKGPAPFYPLEDGTAGEQLHKAMKRYALVPGTIAFTD
                     AHIEVNITYAEYFEMSVRLAEAMKRYGLNTNHRIVVCSENSLQFFMPVLGALFIGVAV
                     APANDIYNERELLNSMNISQPTVVFVSKKGLQKILNVQKKLPIIQKIIIMDSKTDYQG
                     FQSMYTFVTSHLPPGFNEYDFVPESFDRDKTIALIMNSSGSTGLPKGVALPHRTACVR
                     FSHARDPIFGNQIIPDTAILSVVPFHHGFGMFTTLGYLICGFRVVLMYRFEEELFLRS
                     LQDYKIQSALLVPTLFSFFAKSTLIDKYDLSNLHEIASGGAPLSKEVGEAVAKRFHLP
                     GIRQGYGLTETTSAILITPEGDDKPGAVGKVVPFFEAKVVDLDTGKTLGVNQRGELCV
                     RGPMIMSGYVNNPEATNALIDKDGWLHSGDIAYWDEDEHFFIVDRLKSLIKYKGYQVA
                     PAELESILLQHPNIFDAGVAGLPDDDAGELPAAVVVLEHGKTMTEKEIVDYVASQVTT
                     AKKLRGGVVFVDEVPKGLTGKLDARKIREILIKAKKGGKSKL"
misc_feature         3359. .3464
                     /note="Heliothis virescens p63 chaperonin gene double
                     poly-adenylation sequence"
CDS                  4381. .5241
                     /codon_start=1
                     /product="beta-lactamase"
                     /protein_id="AAL30795.1"
                     /db_xref="GI:16904178"
                     /translation="MSIQHFRVALIPFFAAFCLPVFAHPETLVKVKDAEDQLGARVGY
                     IELDLNSGKILESFRPEERFPMMSTFKVLLCGAVLSRIDAGQEQLGRRIHYSQNDLVE
                     YSPVTEKHLTDGMTVRELCSAAITMSDNTAANLLLTTIGGPKELTAFLHNMGDHVTRL
                     DRWEPELNEAIPNDERDTTMPVAMATTLRKLLTGELLTLASRQQLIDWMEADKVAGPL
                     LRSALPAGWFIADKSGAGERGSRGIIAALGPDGKPSRIVVIYTTGSQATMDERNRQIA
                     EIGASLIKHW"
rep_origin           5529. .5948
                     /note="ColE1"
ORIGIN
```

Query Match 81.3%; Score 574.6; DB 12; Length 6182;
Best Local Similarity 99.3%; Pred. No. 1.7e-145;
Matches 577; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy        45 ATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGC 104
             ||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3721 ATCCAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGC 3780

Qy       105 TGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCG 164
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3781 TGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCG 3840

Qy       165 GCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACG 224
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3841 GCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACG 3900

Qy       225 TTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTT 284
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3901 TTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTT 3960

Qy       285 CAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGC 344
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3961 CAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGC 4020

Qy       345 CCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTG 404
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4021 CCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTG 4080

Qy       405 CAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGC 464
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4081 CAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGC 4140

Qy       465 TGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTC 524
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4141 TGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTC 4200

Qy       525 ATTCACGTTTTTGAACCCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGC 584
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4201 ATTCACGTTTTTGAACCCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGC 4260

Qy       585 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCAGAACAC 625
             |||||||||||||||||||||||||||||||||||| | ||
Db      4261 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCATGAAAC 4301
```

```
RESULT 13
AF434932
LOCUS       AF434932                6182 bp    DNA     circular SYN 04-APR-2002
DEFINITION  Expression vector 410-REV, complete sequence.
ACCESSION   AF434932
VERSION     AF434932.1  GI:16904179
KEYWORDS    .
SOURCE      Expression vector 410-REV
  ORGANISM  Expression vector 410-REV
            artificial sequences; vectors.
REFERENCE   1  (bases 1 to 6182)
  AUTHORS   Zieler,H. and Huynh,C.Q.
  TITLE     Intron-dependent stimulation of marker gene expression in cultured
            insect cells
  JOURNAL   Insect Mol. Biol. 11 (1), 87-95 (2002)
  MEDLINE   21830665
   PUBMED   11841506
REFERENCE   2  (bases 1 to 6182)
  AUTHORS   Zieler,H. and Huynh,C.Q.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-OCT-2001) Medical Entomology Section, Laboratory of
            Parasitic Diseases, National Institutes of Health, 4 Center Drive
            MSC 0425, Bethesda, MD 20892-0425, USA
FEATURES             Location/Qualifiers
     source          1. .6182
                     /organism="Expression vector 410-REV"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:176974"
     enhancer        /note="insect-specific"
                     97. .614
                     /note="Autographa californica nuclear polyhedrosis virus
     promoter        hr5 enhancer"
                     615. .1201
                     /note="Autographa californica nuclear polyhedrosis virus
     CDS             immediate early promoter"
                     1328. .2980
                     /codon_start=1
                     /product="firefly luciferase"
                     /protein_id="AAL30796.1"
                     /db_xref="GI:16904180"
                     /translation="MEDAKNIKKGPAPFYPLEDGTAGEQLHKAMKRYALVPGTIAFTD
                     AHIEVNITYAEYFEMSVRLAEAMKRYGLNTNHRIVVCSENSLQFFMPVLGALFIGVAV
                     APANDIYNERELLNSMNISQPTVVFVSKKGLQKILNVQKKLPIIQKIIIMDSKTDYQG
                     FQSMYTFVTSHLPPGFNEYDFVPESFDRDKTIALIMNSSGSTGLPKGVALPHRTACVR
                     FSHARDPIFGNQIIPDTAILSVVPFHHGFGMFTTLGYLICGFRVVLMYRFEEELFLRS
                     LQDYKIQSALLVPTLFSFFAKSTLIDKYDLSNLHEIASGGAPLSKEVGEAVAKRFHLP
                     GIRQGYGLTETTSAILITPEGDDKPGAVGKVVPFFEAKVVDLDTGKTLGVNQRGELCV
                     RGPMIMSGYVNNPEATNALIDKDGWLHSGDIAYWDEDEHFFIVDRLKSLIKYKGYQVA
                     PAELESILLQHPNIFDAGVAGLPDDDAGELPAAVVVLEHGKTMTEKEIVDYVASQVTT
     intron          AKKLRGGVVFVDEVPKGLTGKLDARKIREILIKAKKGGKSKL"
                     complement(3166. .3298)
     misc_feature    /note="artificial intron"
                     3359. .3464
                     /note="Heliothis virescens p63 chaperonin gene double
     CDS             poly-adenylation sequence"
                     4381. .5241
                     /codon_start=1
                     /product="beta-lactamase"
                     /protein_id="AAL30797.1"
                     /db_xref="GI:16904181"
                     /translation="MSIQHFRVALIPFFAAFCLPVFAHPETLVKVKDAEDQLGARVGY
                     IELDLNSGKILESFRPEERFPMMSTFKVLLCGAVLSRIDAGQEQLGRRIHYSQNDLVE
                     YSPVTEKHLTDGMTVRELCSAAITMSDNTAANLLLTTIGGPKELTAFLHNMGDHVTRL
                     DRWEPELNEAIPNDERDTTMPVAMATTLRKLLTGELLTLASRQQLIDWMEADKVAGPL
                     LRSALPAGWFIADKSGAGERGSRGIIAALGPDGKPSRIVVIYTTGSQATMDERNRQIA
     rep_origin      EIGASLIKHW"
                     5529. .5948
                     /note="ColE1"
ORIGIN
```

Query Match 81.3%; Score 574.6; DB 12; Length 6182;
Best Local Similarity 99.3%; Pred. No. 1.7e-145;

```
  Matches  577; Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy         45 ATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGC 104
              ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3721 ATCCAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGC 3780

Qy        105 TGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCG 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3781 TGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCG 3840

Qy        165 GCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACG 224
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3841 GCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACG 3900

Qy        225 TTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTT 284
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3901 TTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTT 3960

Qy        285 CAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGC 344
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3961 CAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGC 4020

Qy        345 CCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTG 404
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4021 CCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTG 4080

Qy        405 CAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGC 464
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4081 CAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGC 4140

Qy        465 TGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTC 524
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4141 TGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTC 4200

Qy        525 ATTCACGTTTTTGAACCCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGC 584
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4201 ATTCACGTTTTTGAACCCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGC 4260

Qy        585 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCAGAACAC 625
              |||||||||||||||||||||||||||||||||||  | ||
Db       4261 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCATGAAAC 4301


RESULT 14
AF434924
LOCUS       AF434924               7421 bp    DNA     circular SYN 04-APR-2002
DEFINITION  Expression vector pACTIN-LUC, complete sequence.
ACCESSION   AF434924
VERSION     AF434924.1  GI:16904155
KEYWORDS    .
SOURCE      Expression vector pACTIN-LUC
  ORGANISM  Expression vector pACTIN-LUC
            artificial sequences; vectors.
REFERENCE   1  (bases 1 to 7421)
  AUTHORS   Zieler,H. and Huynh,C.Q.
  TITLE     Intron-dependent stimulation of marker gene expression in cultured
            insect cells
  JOURNAL   Insect Mol. Biol. 11 (1), 87-95 (2002)
  MEDLINE   21830665
   PUBMED   11841506
REFERENCE   2  (bases 1 to 7421)
  AUTHORS   Zieler,H. and Huynh,C.Q.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-OCT-2001) Medical Entomology Section, Laboratory of
            Parasitic Diseases, National Institutes of Health, 4 Center Drive
            MSC 0425, Bethesda, MD 20892-0425, USA
FEATURES             Location/Qualifiers
     source          1. .7421
                     /organism="Expression vector pACTIN-LUC"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:176966"
                     /note="insect-specific"
     promoter        103. .2660
                     /note="Drosophila melanogaster actin5C promoter"
     CDS             2780. .4432
                     /codon_start=1
                     /product="firefly luciferase"
                     /protein_id="AAL30780.1"
                     /db_xref="GI:16904156"
                     /translation="MEDAKNIKKGPAPFYPLEDGTAGEQLHKAMKRYALVPGTIAFTD
                     AHIEVNITYAEYFEMSVRLAEAMKRYGLNTNHRIVVCSENSLQFFMPVLGALFIGVAV
                     APANDIYNERELLNSMNISQPTVVFVSKKGLQKILNVQKKLPIIQKIIIMDSKTDYQG
                     FQSMYTFVTSHLPPGFNEYDFVPESFDRDKTIALIMNSSGSTGLPKGVALPHRTACVR
                     FSHARDPIFGNQIIPDTAILSVVPFHHGFGMFTTLGYLICGFRVVLMYRFEEELFLRS
                     LQDYKIQSALLVPTLFSFFAKSTLIDKYDLSNLHEIASGGAPLSKEVGEAVAKRFHLP
                     GIRQGYGLTETTSAILITPEGDDKPGAVGKVVPFFEAKVVDLDTGKTLGVNQRGELCV
                     RGPMIMSGYVNNPEATNALIDKDGWLHSGDIAYWDEDEHFFIVDRLKSLIKYKGYQVA
                     PAELESILLQHPNIFDAGVAGLPDDDAGELPAAVVVLEHGKTMTEKEIVDYVASQVTT
                     AKKLRGGVVFVDEVPKGLTGKLDARKIREILIKAKKGGKSKL"
     misc_feature    4598. .4703
                     /note="Heliothis virescens p63 chaperonin gene double
                     poly-adenylation sequence"
     CDS             5620. .6480
                     /codon_start=1
                     /product="beta-lactamase"
                     /protein_id="AAL30781.1"
                     /db_xref="GI:16904157"
                     /translation="MSIQHFRVALIPFFAAFCLPVFAHPETLVKVKDAEDQLGARVGY
                     IELDLNSGKILESFRPEERFPMMSTFKVLLCGAVLSRIDAGQEQLGRRIHYSQNDLVE
                     YSPVTEKHLTDGMTVRELCSAAITMSDNTAANLLLTTIGGPKELTAFLHNMGDHVTRL
                     DRWEPELNEAIPNDERDTTMPVAMATTLRKLLTGELLTLASRQQLIDWMEADKVAGPL
                     LRSALPAGWFIADKSGAGERGSRGIIAALGPDGKPSRIVVIYTTGSQATMDERNRQIA
                     EIGASLIKHW"
     rep_origin      6768. .7187
                     /note="ColE1"
ORIGIN

  Query Match             81.3%;  Score 574.6;  DB 12;  Length 7421;
  Best Local Similarity   99.3%;  Pred. No. 1.8e-145;
  Matches  577;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy         45 ATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGC 104
              ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4960 ATCCAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGC 5019

Qy        105 TGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCG 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5020 TGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCG 5079

Qy        165 GCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACG 224
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5080 GCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACG 5139

Qy        225 TTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTT 284
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5140 TTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTT 5199

Qy        285 CAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGC 344
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5200 CAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGC 5259

Qy        345 CCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTG 404
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5260 CCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTG 5319

Qy        405 CAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGC 464
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5320 CAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGC 5379

Qy        465 TGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTC 524
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5380 TGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTC 5439

Qy        525 ATTCACGTTTTTGAACCCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGC 584
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5440 ATTCACGTTTTTGAACCCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGC 5499
```

```
Qy        585 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCAGAACAC 625
              ||||||||||||||||||||||||||||||||||||| | |||
Db       5500 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCATGAAAC 5540

RESULT 15
AF434928
LOCUS       AF434928                7634 bp    DNA     circular SYN 04-APR-2002
DEFINITION  Expression vector 411-FOR, complete sequence.
ACCESSION   AF434928
VERSION     AF434928.1  GI:16904167
KEYWORDS    .
SOURCE      Expression vector 411-FOR
  ORGANISM  Expression vector 411-FOR
            artificial sequences; vectors.
REFERENCE   1  (bases 1 to 7634)
  AUTHORS   Zieler,H. and Huynh,C.Q.
  TITLE     Intron-dependent stimulation of marker gene expression in cultured
            insect cells
  JOURNAL   Insect Mol. Biol. 11 (1), 87-95 (2002)
  MEDLINE   21830665
   PUBMED   11841506
REFERENCE   2  (bases 1 to 7634)
  AUTHORS   Zieler,H. and Huynh,C.Q.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-OCT-2001) Medical Entomology Section, Laboratory of
            Parasitic Diseases, National Institutes of Health, 4 Center Drive
            MSC 0425, Bethesda, MD 20892-0425, USA
FEATURES             Location/Qualifiers
     source          1. .7634
                     /organism="Expression vector 411-FOR"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:176970"
                     /note="insect-specific"
     promoter        103. .2660
     intron          /note="Drosophila melanogaster actin5C promoter"
                     2718. .2850
     CDS             /note="artificial intron"
                     2993. .4645
                     /codon_start=1
                     /product="firefly luciferase"
                     /protein_id="AAL30788.1"
                     /db_xref="GI:16904168"
                     /translation="MEDAKNIKKGPAPFYPLEDGTAGEQLHKAMKRYALVPGTIAFTD
                     AHIEVNITYAEYFEMSVRLAEAMKRYGLNTNHRIVVCSENSLQFFMPVLGALFIGVAV
                     APANDIYNERELLNSMNISQPTVVFVSKKGLQKILNVQKKLPIIQKIIIMDSKTDYQG
                     FQSMYTFVTSHLPPGFNEYDFVPESFDRDKTIALIMNSSGSTGLPKGVALPHRTACVR
                     FSHARDPIFGNQIIPDTAILSVVPFHHGFGMFTTLGYLICGFRVVLMYRFEEELFLRS
                     LQDYKIQSALLVPTLFSFFAKSTLIDKYDLSNLHEIASGGAPLSKEVGEAVAKRFHLP
                     GIRQGYGLTETTSAILITPEGDDKPGAVGKVVPFFEAKVVDLDTGKTLGVNQRGELCV
                     RGPMIMSGYVNNPEATNALIDKDGWLHSGDIAYWDEDEHFFIVDRLKSLIKYKGYQVA
                     PAELESILLQHPNIFDAGVAGLPDDDAGELPAAVVVLEHGKTMTEKEIVDYVASQVTT
                     AKKLRGGVVFVDEVPKGLTGKLDARKIREILIKAKKGGKSKL"
     misc_feature    4811. .4916
                     /note="Heliothis virescens p63 chaperonin gene double
                     poly-adenylation sequence"
     CDS             5833. .6693
                     /codon_start=1
                     /product="beta-lactamase"
                     /protein_id="AAL30789.1"
                     /db_xref="GI:16904169"
                     /translation="MSIQHFRVALIPFFAAFCLPVFAHPETLVKVKDAEDQLGARVGY
                     IELDLNSGKILESFRPEERFPMMSTFKVLLCGAVLSRIDAGQEQLGRRIHYSQNDLVE
                     YSPVTEKHLTDGMTVRELCSAAITMSDNTAANLLLTTIGGPKELTAFLHNMGDHVTRL
                     DRWEPELNEAIPNDERDTTMPVAMATTLRKLLTGELLTLASRQQLIDWMEADKVAGPL
                     LRSALPAGWFIADKSGAGERGSRGIIAALGPDGKPSRIVVIYTTGSQATMDERNRQIA
                     EIGASLIKHW"
     rep_origin      6981. .7400
                     /note="ColE1"
ORIGIN

  Query Match             81.3%;  Score 574.6;  DB 12;  Length 7634;
```

```
  Best Local Similarity   99.3%;  Pred. No. 1.8e-145;
  Matches  577;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy         45 ATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGC 104
              ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5173 ATCCAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGC 5232

Qy        105 TGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCG 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5233 TGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCG 5292

Qy        165 GCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACG 224
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5293 GCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACG 5352

Qy        225 TTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTT 284
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5353 TTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTT 5412

Qy        285 CAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGC 344
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5413 CAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGC 5472

Qy        345 CCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTG 404
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5473 CCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTG 5532

Qy        405 CAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGC 464
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5533 CAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGC 5592

Qy        465 TGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTC 524
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5593 TGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTC 5652

Qy        525 ATTCACGTTTTTGAACCCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGC 584
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5653 ATTCACGTTTTTGAACCCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGC 5712

Qy        585 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCAGAACAC 625
              ||||||||||||||||||||||||||||||||||||| | |||
Db       5713 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCATGAAAC 5753

Search completed: January 26, 2005, 08:24:53
Job time : 3017.51 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: January 26, 2005, 02:11:40 ; Search time 2221.98 Seconds
(without alignments)
11594.581 Million cell updates/sec

Title: US-10-826-523-40
Perfect score: 707
Sequence: 1 ggatcccatgcgtcaatttt.....aaaattgacgcatgggatcc 707

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_htc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
c 1	585.4	82.8	682	7	CK781302	CK781302 UI-M-GI0-
c 2	585.4	82.8	703	7	CF537771	CF537771 UI-M-GI0-
c 3	585.4	82.8	711	7	CF743678	CF743678 UI-M-GI0-
4	585.4	82.8	853	9	CR065566	CR065566 Forward s
c 5	584.4	82.7	703	5	BQ154655	BQ154655 NF087C06I
c 6	584.4	82.7	719	6	CD350897	CD350897 UI-M-GI0-
c 7	583.8	82.6	751	9	CR087413	CR087413 Forward s
c 8	580.4	82.1	716	9	CR131675	CR131675 Forward s
c 9	580	82.0	786	9	CR108810	CR108810 Forward s
c 10	565.4	80.0	691	6	CD350776	CD350776 UI-M-GI0-
11	564	79.8	762	9	CR035207	CR035207 Reverse s
c 12	563.4	79.7	749	6	CD351273	CD351273 UI-M-GI0-
c 13	562	79.5	733	9	CR139473	CR139473 Forward s
c 14	555	78.5	633	9	CR167678	CR167678 Reverse s
c 15	554.4	78.4	791	6	CB520716	CB520716 UI-M-GI0-
c 16	550.4	77.9	630	9	CR013093	CR013093 Reverse s
17	549	77.7	702	9	CR084987	CR084987 Reverse s
18	545.4	77.1	611	4	BI423083	BI423083 EST533749
c 19	543.4	76.9	716	7	CK781284	CK781284 UI-M-GI0-
c 20	540.8	76.5	716	7	CF851544	CF851544 psMA007xC
c 21	537.2	76.0	706	1	AV731514	AV731514 AV731514
22	535.4	75.7	730	9	CR077673	CR077673 Reverse s
c 23	525	74.3	635	9	BX982794	BX982794 Forward s
24	521.8	73.8	536	9	CR026633	CR026633 Reverse s
c 25	516	73.0	887	9	CR192748	CR192748 Forward s
c 26	515	72.8	885	9	CR255010	CR255010 Forward s
c 27	498.4	70.5	683	6	CD348896	CD348896 UI-M-FY0-
28	493	69.7	495	4	BG457760	BG457760 NF036E03P
c 29	490.4	69.4	786	6	CB520098	CB520098 UI-M-GI0-
c 30	482.6	68.3	795	6	CD351738	CD351738 UI-M-GI0-
c 31	477.4	67.5	760	7	CF538326	CF538326 UI-M-GI0-
32	474.4	67.1	620	9	CR034585	CR034585 Forward s
c 33	474.4	67.1	620	9	CR086282	CR086282 Reverse s
c 34	474.4	67.1	620	9	CR171547	CR171547 Reverse s
35	471.8	66.7	620	9	CR078862	CR078862 Forward s
c 36	467.4	66.1	752	9	CR088896	CR088896 Forward s
c 37	459.6	65.0	805	6	CD352530	CD352530 UI-M-GI0-
c 38	457	64.6	676	5	BQ154673	BQ154673 NF095C11I
39	455.4	64.4	483	9	CR068828	CR068828 Reverse s
c 40	447.4	63.3	547	9	CR148906	CR148906 Forward s
c 41	447.4	63.3	885	9	CR024214	CR024214 Reverse s
42	444.4	62.9	599	9	CR079016	CR079016 Forward s
43	443.4	62.7	702	9	CR038019	CR038019 Forward s
c 44	440.8	62.3	538	9	CR029330	CR029330 Forward s
c 45	437.6	61.9	901	6	CA328434	CA328434 UI-M-FY0-

ALIGNMENTS

RESULT 1
CK781302/c
LOCUS CK781302 682 bp mRNA linear EST 23-FEB-2004
DEFINITION UI-M-GI0-clp-g-05-0-UI.r1 NIH_BMAP_GI0 Mus musculus cDNA clone IMAGE:30619060 5', mRNA sequence.
ACCESSION CK781302
VERSION CK781302.1 GI:42746980
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 682)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. James Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at http://genome.uiowa.edu/distribution/mousefl.html
This clone was contributed by the Brain Molecular Anatomy Project (BMAP)
Seq primer: pYX-5.
FEATURES Location/Qualifiers
source 1. .682
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:30619060"
/tissue_type="whole brain"
/dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_GI0"
/note="Organ: Brain; Vector: pYX- Asc; Site_1: EcoR I; Site_2: Not I; The library was constructed according Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel.First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site .Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector . The library tag

sequence located between the Not I site and the polyA tail is AGCGAGACAG. This library was created for the University Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN

Query Match 82.8%; Score 585.4; DB 7; Length 682;
Best Local Similarity 99.8%; Pred. No. 2.3e-164;
Matches 586; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy      45 ATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGC 104
           ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     611 ATCCAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGC 552

Qy     105 TGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCG 164
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     551 TGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCG 492

Qy     165 GCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACG 224
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     491 GCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACG 432

Qy     225 TTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTT 284
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     431 TTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTT 372

Qy     285 CAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGC 344
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     371 CAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGC 312

Qy     345 CCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTG 404
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     311 CCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTG 252

Qy     405 CAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGC 464
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     251 CAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGC 192

Qy     465 TGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTC 524
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     191 TGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTC 132

Qy     525 ATTCACGTTTTTGAACCCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGC 584
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     131 ATTCACGTTTTTGAACCCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGC 72

Qy     585 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCAGAACACGCAGCT 631
           |||||||||||||||||||||||||||||||||||||||||||||||
Db      71 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCAGAACACGCAGCT 25
```

RESULT 2
CF537771/c
LOCUS CF537771 703 bp mRNA linear EST 12-SEP-2003
DEFINITION UI-M-GI0-chn-m-08-0-UI.r1 NIH_BMAP_GI0 Mus musculus cDNA clone IMAGE:30537799 5', mRNA sequence.
ACCESSION CF537771
VERSION CF537771.1 GI:34589753
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 703)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at http://genome.uiowa.edu/distribution/mousefl.html
This clone was contributed by the Brain Molecular Anatomy Project (BMAP)
Seq primer: pYX-5.

FEATURES Location/Qualifiers
source 1. .703
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:30537799"
/tissue_type="whole brain"
/dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_GI0"
/note="Organ: Brain; Vector: pYX- Asc; Site_1: EcoR I; Site_2: Not I; The library was constructed according Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel.First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site .Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector . The library tag sequence located between the Not I site and the polyA tail is AGCGAGACAG. This library was created for the University Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN

Query Match 82.8%; Score 585.4; DB 7; Length 703;
Best Local Similarity 99.8%; Pred. No. 2.3e-164;
Matches 586; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy      45 ATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGC 104
           ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     643 ATCCAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGC 584

Qy     105 TGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCG 164
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     583 TGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCG 524

Qy     165 GCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACG 224
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     523 GCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACG 464

Qy     225 TTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTT 284
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     463 TTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTT 404

Qy     285 CAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGC 344
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     403 CAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGC 344

Qy     345 CCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTG 404
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     343 CCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTG 284

Qy     405 CAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGC 464
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     283 CAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGC 224

Qy     465 TGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTC 524
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     223 TGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTC 164

Qy     525 ATTCACGTTTTTGAACCCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGC 584
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        163 ATTCACGTTTTTGAACCCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGC 104

Qy        585 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCAGAACACGCAGCT 631
              |||||||||||||||||||||||||||||||||||||||||||||||
Db        103 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCAGAACACGCAGCT 57


RESULT 3
CF743678/c
LOCUS       CF743678                711 bp    mRNA    linear   EST 10-OCT-2003
DEFINITION  UI-M-GI0-clq-e-19-0-UI.r1 NIH_BMAP_GI0 Mus musculus cDNA clone
            IMAGE:30616722 5', mRNA sequence.
ACCESSION   CF743678
VERSION     CF743678.1  GI:37640017
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 711)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. James Lin, University of Iowa
             cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
             Clone Distribution: Distribution information can be found at
            http://genome.uiowa.edu/distribution/mousefl.html
             This clone was contributed by the Brain Molecular Anatomy Project
            (BMAP)
            Seq primer: pYX-5.
FEATURES             Location/Qualifiers
     source          1. .711
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:30616722"
                     /tissue_type="whole brain"
                     /dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
                     /lab_host="DH10B (T1 phage resistant)"
                     /clone_lib="NIH_BMAP_GI0"
                     /note="Organ: Brain; Vector: pYX- Asc; Site_1: EcoR I;
                     Site_2: Not I; The library was constructed according
                     Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
                     1996. Denatured RNA was size fractionated on a 1% agarose
                     gel.First strand cDNA synthesis was primed with oligo-dT
                     primer containing a Not I site .Double strand cDNA was
                     size selected according to mRNA size fraction, ligated
                     with EcoR I adaptor, digested with NotI and then cloned
                     directionally into pYX-Asc vector . The library tag
                     sequence located between the Not I site and the polyA tail
                     is AGCGAGACAG. This library was created for the University
                     Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
                     Developing Mouse Nervous System', supported by National
                     Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
                     program coordinator."
ORIGIN

  Query Match             82.8%;  Score 585.4;  DB 7;  Length 711;
  Best Local Similarity   99.8%;  Pred. No. 2.3e-164;
  Matches  586;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         45 ATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGC 104
              ||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        654 ATCCAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGC 595

Qy        105 TGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCG 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        594 TGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCG 535

Qy        165 GCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACG 224
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        534 GCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACG 475

Qy        225 TTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTT 284
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        474 TTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTT 415

Qy        285 CAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGC 344
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        414 CAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGC 355

Qy        345 CCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTG 404
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        354 CCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTG 295

Qy        405 CAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGC 464
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        294 CAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGC 235

Qy        465 TGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTC 524
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        234 TGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTC 175

Qy        525 ATTCACGTTTTTGAACCCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGC 584
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        174 ATTCACGTTTTTGAACCCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGC 115

Qy        585 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCAGAACACGCAGCT 631
              |||||||||||||||||||||||||||||||||||||||||||||||
Db        114 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCAGAACACGCAGCT 68


RESULT 4
CR065566
LOCUS       CR065566                853 bp    DNA     linear   GSS 05-JUL-2004
DEFINITION  Forward strand read from insert in 3'HPRT insertion targeting and
            chromosome engineering clone MHPP250k12, genomic survey sequence.
ACCESSION   CR065566
VERSION     CR065566.1  GI:49799156
KEYWORDS    GSS; genome survey sequence; MICER.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 853)
  AUTHORS   Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L.,
            Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,
            Rogers,J. and Bradley,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. http://www.sanger.ac.uk/MICER
FEATURES             Location/Qualifiers
     source          1. .853
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /clone="MHPP250k12"
                     /clone_lib="MHPP"
ORIGIN

  Query Match             82.8%;  Score 585.4;  DB 9;  Length 853;
  Best Local Similarity   99.8%;  Pred. No. 2.4e-164;
  Matches  586;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         45 ATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGC 104
              ||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        264 ATCCAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGC 323

Qy        105 TGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCG 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        324 TGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCG 383
```

```
Qy        165 GCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACG 224
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        384 GCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACG 443

Qy        225 TTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTT 284
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        444 TTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTT 503

Qy        285 CAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGC 344
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        504 CAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGC 563

Qy        345 CCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTG 404
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        564 CCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTG 623

Qy        405 CAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGC 464
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        624 CAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGC 683

Qy        465 TGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTC 524
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        684 TGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTC 743

Qy        525 ATTCACGTTTTTGAACCCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGC 584
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        744 ATTCACGTTTTTGAACCCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGC 584

Qy        585 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCAGAACACGCAGCT 631
              ||||||||||||||||||||||||||||||||||||||||||||||
Db        804 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCAGAACACGCAGCT 850
```

RESULT 5
BQ154655/c
LOCUS BQ154655 703 bp mRNA linear EST 24-APR-2002
DEFINITION NF087C06IR1F1050 Irradiated Medicago truncatula cDNA clone NF087C06IR 5', mRNA sequence.
ACCESSION BQ154655
VERSION BQ154655.1 GI:20291714
KEYWORDS EST.
SOURCE Medicago truncatula (barrel medic)
ORGANISM Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.
REFERENCE 1 (bases 1 to 703)
AUTHORS Torres-Jerez,I., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell,C.J., Flores,H.R., Inman,J.T., Weller,J.W. and May,G.D.
TITLE Expressed Sequence Tags from the Samuel Roberts Noble Foundation Medicago truncatula irradiated library
JOURNAL Unpublished (2001)
COMMENT Contact: May GD
Plant Biology Division
The Samuel Roberts Noble Foundation
2510 Sam Noble Parkway, Ardmore, OK 73402, USA
Tel: 580 224 6650
Fax: 580 224 6692
Email: gdmay@noble.org
Insert Length: 703 Std Error: 0.00
Plate: 087 row: C column: 06
Seq primer: TCACACAGGAAACAGCTATGAC.
FEATURES Location/Qualifiers
source 1. .703
/organism="Medicago truncatula"
/mol_type="mRNA"
/db_xref="taxon:3880"
/clone="NF087C06IR"
/tissue_type="seedlings"
/dev_stage="seedling"
/clone_lib="Irradiated"
/note="Vector: Lambda Zap; Seedlings were exposed either to 100 gy gamma or 0.5, 1, 5, or 10 kJ/m2 UV irradiation. Gamma-irradiated samples were harvested at 6, 12, 24 and 48 hours after treatment. UV-irradiated samples were harvested 24 hours post-treatment. cDNA was prepared from polyA+ enriched, pooled samples of equivalent amounts of total RNA from each sample. The cDNA was directionally ligated into the Uni-Zap XR vector (Stratagene) and packaged using the Gigapack III Gold packaging extracts. Phagemids containing cDNA inserts were in vivo excised from the recombinant Uni-ZAP XR vector using ExAssist helper phage and the E. coli strain XL1-Blue MRF' (Stratagene). Excised plasmids were plated using SOLR cells."
ORIGIN

Query Match 82.7%; Score 584.4; DB 5; Length 703;
Best Local Similarity 99.7%; Pred. No. 4.6e-164;
Matches 585; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy         45 ATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGC 104
              ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        624 ATCCAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGC 565

Qy        105 TGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCG 164
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        564 TGGCGCTATNTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCG 505

Qy        165 GCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACG 224
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        504 GCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACG 445

Qy        225 TTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTT 284
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        444 TTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTT 385

Qy        285 CAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGC 344
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        384 CAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGC 325

Qy        345 CCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTG 404
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        324 CCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTG 265

Qy        405 CAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGC 464
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        264 CAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGC 205

Qy        465 TGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTC 524
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        204 TGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTC 145

Qy        525 ATTCACGTTTTTGAACCCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGC 584
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        144 ATTCACGTTTTTGAACCCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGC 85

Qy        585 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCAGAACACGCAGCT 631
              ||||||||||||||||||||||||||||||||||||||||||||||
Db         84 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCAGAACACGCAGCT 38
```

RESULT 6
CD350897/c
LOCUS CD350897 719 bp mRNA linear EST 09-JUL-2003
DEFINITION UI-M-GI0-cgh-a-18-0-UI.r1 NIH_BMAP_GI0 Mus musculus cDNA clone IMAGE:6853243 5', mRNA sequence.
ACCESSION CD350897
VERSION CD350897.1 GI:31142412
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

```
REFERENCE   1  (bases 1 to 719)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Jim Lin, University of Iowa
             cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, University of Iowa,
             Clone Distribution: Distribution information can be found at
            http://genome.uiowa.edu/distribution/mousefl.html
             This clone was contributed by the Brain Molecular Anatomy Project
            (BMAP)
            Seq primer: pYX-5.
FEATURES             Location/Qualifiers
     source          1. .719
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:6853243"
                     /tissue_type="whole brain"
                     /dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
                     /lab_host="DH10B (T1 phage resistant)"
                     /clone_lib="NIH_BMAP_GI0"
                     /note="Organ: Brain; Vector: pYX- Asc; Site_1: EcoR I;
                     Site_2: Not I; The library was constructed according
                     Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
                     1996. Denatured RNA was size fractionated on a 1% agarose
                     gel.First strand cDNA synthesis was primed with oligo-dT
                     primer containing a Not I site .Double strand cDNA was
                     size selected according to mRNA size fraction, ligated
                     with EcoR I adaptor, digested with NotI and then cloned
                     directionally into pYX-Asc vector . The library tag
                     sequence located between the Not I site and the polyA tail
                     is AGCGAGACAG. This library was created for the University
                     Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
                     Developing Mouse Nervous System', supported by National
                     Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
                     program coordinator."
ORIGIN

  Query Match           82.7%;  Score 584.4;  DB 6;  Length 719;
  Best Local Similarity   99.7%;  Pred. No. 4.6e-164;
  Matches  585;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy         45 ATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGC 104
              ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        663 ATCCAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGC 604

Qy        105 TGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCG 164
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        603 TGGCGCTATCTGGGCATCGGNGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCG 544

Qy        165 GCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACG 224
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        543 GCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACG 484

Qy        225 TTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTT 284
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        483 TTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTT 424

Qy        285 CAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGC 344
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        423 CAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGC 364

Qy        345 CCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTG 404
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 CCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTG 304

Qy        405 CAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGC 464
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 CAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGC 244

Qy        465 TGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTC 524
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 TGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTC 184

Qy        525 ATTCACGTTTTTGAACCCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGC 584
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 ATTCACGTTTTTGAACCCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGC 124

Qy        585 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCAGAACACGCAGCT 631
              |||||||||||||||||||||||||||||||||||||||||||||||
Db        123 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCAGAACACGCAGCT 77


RESULT 7
CR087413/c
LOCUS       CR087413                751 bp    DNA     linear   GSS 05-JUL-2004
DEFINITION  Forward strand read from insert in 3'HPRT insertion targeting and
            chromosome engineering clone MHPP438f24, genomic survey sequence.
ACCESSION   CR087413
VERSION     CR087413.1  GI:49821005
KEYWORDS    GSS; genome survey sequence; MICER.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 751)
  AUTHORS   Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L.,
            Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,
            Rogers,J. and Bradley,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. http://www.sanger.ac.uk/MICER
FEATURES             Location/Qualifiers
     source          1. .751
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /clone="MHPP438f24"
                     /clone_lib="MHPP"
ORIGIN

  Query Match           82.6%;  Score 583.8;  DB 9;  Length 751;
  Best Local Similarity   99.7%;  Pred. No. 7e-164;
  Matches  585;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy         45 ATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGC 104
              ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        623 ATCCAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGC 564

Qy        105 TGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCG 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        563 TGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCG 504

Qy        165 GCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACG 224
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        503 GCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACG 444

Qy        225 TTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTT 284
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        443 TTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTT 384

Qy        285 CAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGC 344
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        383 CAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGC 324

Qy        345 CCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTG 404
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        323 CCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTG 264

Qy        405 CAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGC 464
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        263 CAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGC 204
Qy        465 TGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTC 524
Db        203 TGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTC 524
Qy        525 ATTCACGTTTTTGAACCCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGC 144
Db        143 ATTCACGTTTTTGAACCCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGC 584
Qy        585 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCAGAACACGCAGCT 631
Db         83 GTGATGGAGCAGATGAAGATGCTCGACACGCAGCAGAACACGCAGCT 37

RESULT 8
CR131675/c
LOCUS       CR131675                716 bp    DNA     linear   GSS 06-JUL-2004
DEFINITION  Forward strand read from insert in 3'HPRT insertion targeting and
            chromosome engineering clone MHPP438f23, genomic survey sequence.
ACCESSION   CR131675
VERSION     CR131675.1  GI:49879128
KEYWORDS    GSS; genome survey sequence; MICER.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 716)
  AUTHORS   Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L.,
            Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,
            Rogers,J. and Bradley,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. http://www.sanger.ac.uk/MICER
FEATURES             Location/Qualifiers
     source          1. .716
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /clone="MHPP438f23"
                     /clone_lib="MHPP"
ORIGIN

  Query Match             82.1%;  Score 580.4;  DB 9;  Length 716;
  Best Local Similarity   99.8%;  Pred. No. 7.3e-163;
  Matches  581;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;
Qy         45 ATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGC 104
Db        582 ATCCAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGC 104
Qy        105 TGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCG 164
Db        522 TGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCG 164
Qy        165 GCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACG 224
Db        462 GCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACG 224
Qy        225 TTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTT 284
Db        402 TTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTT 284
Qy        285 CAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGC 344
Db        342 CAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGC 344
Qy        345 CCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTG 404
Db        282 CCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTG 404
Qy        405 CAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGC 464
Db        222 CAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGC 163
Qy        465 TGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTC 524
Db        162 TGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTC 524
Qy        525 ATTCACGTTTTTGAACCCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGC 584
Db        102 ATTCACGTTTTTGAACCCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGC 584
Qy        585 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCAGAACACG 626
Db         42 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCAGAACACG 1

RESULT 9
CR108810/c
LOCUS       CR108810                786 bp    DNA     linear   GSS 05-JUL-2004
DEFINITION  Forward strand read from insert in 3'HPRT insertion targeting and
            chromosome engineering clone MHPP22o16, genomic survey sequence.
ACCESSION   CR108810
VERSION     CR108810.1  GI:49856225
KEYWORDS    GSS; genome survey sequence; MICER.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 786)
  AUTHORS   Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L.,
            Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,
            Rogers,J. and Bradley,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. http://www.sanger.ac.uk/MICER
FEATURES             Location/Qualifiers
     source          1. .786
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /clone="MHPP22o16"
                     /clone_lib="MHPP"
ORIGIN

  Query Match             82.0%;  Score 580;  DB 9;  Length 786;
  Best Local Similarity   100.0%;  Pred. No. 9.8e-163;
  Matches  580;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
Qy         52 TGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGCTGGCGCT 111
Db        786 TGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGCTGGCGCT 111
Qy        112 ATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCGGCATGGA 171
Db        726 ATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCGGCATGGA 727
Qy        172 AAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACGTTTACCA 231
Db        666 AAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACGTTTACCA 667
Qy        232 TGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTTCAGGCCA 291
Db        606 TGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTTCAGGCCA 607
Qy        292 CCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGCCCGAAGC 351
Db        546 CCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGCCCGAAGC 547
Qy        352 GCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTGCAGATTA 411
Db        486 GCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTGCAGATTA 487
Qy        412 ATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGCTGGATGC 471
```

```
Db        426 ATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGCTGGATGC 367

Qy        472 CGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTCATTCACG 531
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        366 CGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTCATTCACG 307

Qy        532 TTTTTGAACCCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGCGTGATGG 591
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        306 TTTTTGAACCCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGCGTGATGG 247

Qy        592 AGCAGATGAAGATGCTCGACACGCTGCAGAACACGCAGCT 631
              |||||||||||||||||||||||||||||||||||||||
Db        246 AGCAGATGAAGATGCTCGACACGCTGCAGAACACGCAGCT 207


RESULT 10
CD350776/c
LOCUS       CD350776                691 bp    mRNA    linear   EST 09-JUL-2003
DEFINITION  UI-M-GI0-cgh-g-19-0-UI.r1 NIH_BMAP_GI0 Mus musculus cDNA clone
            IMAGE:6853388 5', mRNA sequence.
ACCESSION   CD350776
VERSION     CD350776.1  GI:31142363
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 691)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Jim Lin, University of Iowa
             cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
             Clone Distribution: Distribution information can be found at
            http://genome.uiowa.edu/distribution/mousefl.html
             This clone was contributed by the Brain Molecular Anatomy Project
            (BMAP)
            Seq primer: pYX-5.
FEATURES             Location/Qualifiers
     source          1. .691
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:6853388"
                     /tissue_type="whole brain"
                     /dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
                     /lab_host="DH10B (T1 phage resistant)"
                     /clone_lib="NIH_BMAP_GI0"
                     /note="Organ: Brain; Vector: pYX- Asc; Site_1: EcoR I;
                     Site_2: Not I; The library was constructed according
                     Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
                     1996. Denatured RNA was size fractionated on a 1% agarose
                     gel.First strand cDNA synthesis was primed with oligo-dT
                     primer containing a Not I site .Double strand cDNA was
                     size selected according to mRNA size fraction, ligated
                     with EcoR I adaptor, digested with NotI and then cloned
                     directionally into pYX-Asc vector . The library tag
                     sequence located between the Not I site and the polyA tail
                     is AGCGAGACAG. This library was created for the University
                     Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
                     Developing Mouse Nervous System', supported by National
                     Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
                     program coordinator."
ORIGIN

  Query Match            80.0%;  Score 565.4;  DB 6;  Length 691;
  Best Local Similarity  99.8%;  Pred. No. 2.3e-158;
  Matches  566;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         45 ATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGC 104
              ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        567 ATCCAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGC 508

Qy        105 TGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCG 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        507 TGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCG 448

Qy        165 GCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACG 224
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        447 GCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACG 388

Qy        225 TTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTT 284
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        387 TTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTT 328

Qy        285 CAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGC 344
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        327 CAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGC 268

Qy        345 CCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTG 404
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        267 CCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTG 208

Qy        405 CAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGC 464
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        207 CAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGC 148

Qy        465 TGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTC 524
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        147 TGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTC 88

Qy        525 ATTCACGTTTTTGAACCCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGC 584
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         87 ATTCACGTTTTTGAACCCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGC 28

Qy        585 GTGATGGAGCAGATGAAGATGCTCGAC 611
              |||||||||||||||||||||||||||
Db         27 GTGATGGAGCAGATGAAGATGCTCGAC 1


RESULT 11
CR035207
LOCUS       CR035207                762 bp    DNA     linear   GSS 05-JUL-2004
DEFINITION  Reverse strand read from insert in 5'HPRT insertion targeting and
            chromosome engineering clone MHPN278n13, genomic survey sequence.
ACCESSION   CR035207
VERSION     CR035207.1  GI:49768262
KEYWORDS    GSS; genome survey sequence; MICER.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 762)
  AUTHORS   Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L.,
            Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,
            Rogers,J. and Bradley,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. http://www.sanger.ac.uk/MICER
FEATURES             Location/Qualifiers
     source          1. .762
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /clone="MHPN278n13"
                     /clone_lib="MHPN"
ORIGIN

  Query Match            79.8%;  Score 564;  DB 9;  Length 762;
  Best Local Similarity  100.0%;  Pred. No. 6.2e-158;
  Matches  564;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy         68 CGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGGCATCGGGGA 127
Db          2 CGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGGCATCGGGGA 61

Qy        128 GGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCGGCATGGAAAGAGTTTGCCGAGGA 187
Db         62 GGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCGGCATGGAAAGAGTTTGCCGAGGA 121

Qy        188 TGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACGTTTACCATGATGATTCGGGAAGG 247
Db        122 TGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACGTTTACCATGATGATTCGGGAAGG 181

Qy        248 TGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTTCAGGCCACCTGGGATACCAGTTC 307
Db        182 TGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTTCAGGCCACCTGGGATACCAGTTC 241

Qy        308 GTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGCCCGAAGCGCATCAGCAACCCGAA 367
Db        242 GTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGCCCGAAGCGCATCAGCAACCCGAA 301

Qy        368 CAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTGCAGATTAATGACAGCGGTGCGGC 427
Db        302 CAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTGCAGATTAATGACAGCGGTGCGGC 361

Qy        428 GCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGCTGGATGCCGCAGAAATGGACATG 487
Db        362 GCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGCTGGATGCCGCAGAAATGGACATG 421

Qy        488 GATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTCATTCACGTTTTTGAACCCGTGGA 547
Db        422 GATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTCATTCACGTTTTTGAACCCGTGGA 481

Qy        548 GGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGCGTGATGGAGCAGATGAAGATGCT 607
Db        482 GGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGCGTGATGGAGCAGATGAAGATGCT 541

Qy        608 CGACACGCTGCAGAACACGCAGCT 631
Db        542 CGACACGCTGCAGAACACGCAGCT 565
```

RESULT 12
CD351273/c
LOCUS CD351273 749 bp mRNA linear EST 09-JUL-2003
DEFINITION UI-M-GI0-cgh-f-22-0-UI.r1 NIH_BMAP_GI0 Mus musculus cDNA clone IMAGE:6853367 5', mRNA sequence.
ACCESSION CD351273
VERSION CD351273.1 GI:31142608
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 749)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at http://genome.uiowa.edu/distribution/mousefl.html
This clone was contributed by the Brain Molecular Anatomy Project (BMAP)
Seq primer: pYX-5.
FEATURES Location/Qualifiers
source 1. .749
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:6853367"
/tissue_type="whole brain"
/dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_GI0"
/note="Organ: Brain; Vector: pYX- Asc; Site_1: EcoR I; Site_2: Not I; The library was constructed according Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel.First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site .Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector . The library tag sequence located between the Not I site and the polyA tail is AGCGAGACAG. This library was created for the University Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."
ORIGIN

Query Match 79.7%; Score 563.4; DB 6; Length 749;
Best Local Similarity 99.8%; Pred. No. 9.4e-158;
Matches 564; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy         45 ATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGC 104
Db        565 ATCCAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGC 506

Qy        105 TGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCG 164
Db        505 TGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCG 446

Qy        165 GCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACG 224
Db        445 GCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACG 386

Qy        225 TTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTT 284
Db        385 TTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTT 326

Qy        285 CAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGC 344
Db        325 CAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGC 266

Qy        345 CCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTG 404
Db        265 CCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTG 206

Qy        405 CAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGC 464
Db        205 CAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGC 146

Qy        465 TGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTC 524
Db        145 TGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTC 86

Qy        525 ATTCACGTTTTTGAACCCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGC 584
Db         85 ATTCACGTTTTTGAACCCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGC 26

Qy        585 GTGATGGAGCAGATGAAGATGCTCG 609
Db         25 GTGATGGAGCAGATGAAGATGCTCG 1
```

RESULT 13
CR139473/c
LOCUS CR139473 733 bp DNA linear GSS 06-JUL-2004
DEFINITION Forward strand read from insert in 3'HPRT insertion targeting and

```
          chromosome engineering clone MHPP274b23, genomic survey sequence.
ACCESSION   CR139473
VERSION     CR139473.1  GI:49887345
KEYWORDS    GSS; genome survey sequence; MICER.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 733)
  AUTHORS   Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L.,
            Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,
            Rogers,J. and Bradley,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. http://www.sanger.ac.uk/MICER
FEATURES             Location/Qualifiers
     source          1. .733
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /clone="MHPP274b23"
                     /clone_lib="MHPP"
ORIGIN

  Query Match             79.5%;  Score 562;  DB 9;  Length 733;
  Best Local Similarity   99.3%;  Pred. No. 2.5e-157;
  Matches  573;  Conservative    0;  Mismatches    3;  Indels    1;  Gaps    1;

Qy         56 TCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGCTGGCGCTATCT 115
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        733 TCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGCTGGCGCTATCT 674

Qy        116 GGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCGGCATGGAAAGA 175
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        673 GGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCGGCATGGAAAGA 614

Qy        176 GTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACGTTTACCATGAT 235
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        613 GTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACGTTTACCATGAT 554

Qy        236 GATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTTCAGGCCACCTG 295
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        553 GATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTTCAGGCCACCTG 494

Qy        296 GGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGCCCGAAGCGCAT 355
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        493 GGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGCCCGAAGCGCAT 434

Qy        356 CAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCG-TGCCGGTGTGCAGATTAATG 414
              ||||||||||||||||||||||||||||||||||||    |||| ||||||||||||||||
Db        433 CAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCNNCTGCCNGTGTGCAGATTAATG 374

Qy        415 ACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGCTGGATGCCGC 474
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        373 ACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGCTGGATGCCGC 314

Qy        475 AGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTCATTCACGTTT 534
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        313 AGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTCATTCACGTTT 254

Qy        535 TTGAACCCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGCGTGATGGAGC 594
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        253 TTGAACCCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGCGTGATGGAGC 194

Qy        595 AGATGAAGATGCTCGACACGCTGCAGAACACGCAGCT 631
              |||||||||||||||||||||||||||||||||||||
Db        193 AGATGAAGATGCTCGACACGCTGCAGAACACGCAGCT 157


RESULT 14
CR167678/c
LOCUS       CR167678                633 bp    DNA     linear   GSS 06-JUL-2004
DEFINITION  Reverse strand read from insert in 3'HPRT insertion targeting and
```

```
          chromosome engineering clone MHPP161k02, genomic survey sequence.
ACCESSION   CR167678
VERSION     CR167678.1  GI:49946527
KEYWORDS    GSS; genome survey sequence; MICER.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 633)
  AUTHORS   Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L.,
            Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,
            Rogers,J. and Bradley,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. http://www.sanger.ac.uk/MICER
FEATURES             Location/Qualifiers
     source          1. .633
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /clone="MHPP161k02"
                     /clone_lib="MHPP"
ORIGIN

  Query Match             78.5%;  Score 555;  DB 9;  Length 633;
  Best Local Similarity   100.0%;  Pred. No. 3e-155;
  Matches  555;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         77 TTTTTTCCGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGGCATCGGGGAGGAAGAAGC 136
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        633 TTTTTTCCGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGGCATCGGGGAGGAAGAAGC 574

Qy        137 CCGTGCCTTTTCCCGCGAGGTTGAAGCGGCATGGAAAGAGTTTGCCGAGGATGACTGCTG 196
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        573 CCGTGCCTTTTCCCGCGAGGTTGAAGCGGCATGGAAAGAGTTTGCCGAGGATGACTGCTG 514

Qy        197 CTGCATTGACGTTGAGCGAAAACGCACGTTTACCATGATGATTCGGGAAGGTGTGGCCAT 256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        513 CTGCATTGACGTTGAGCGAAAACGCACGTTTACCATGATGATTCGGGAAGGTGTGGCCAT 454

Qy        257 GCACGCCTTTAACGGTGAACTGTTCGTTCAGGCCACCTGGGATACCAGTTCGTCGCGGCT 316
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        453 GCACGCCTTTAACGGTGAACTGTTCGTTCAGGCCACCTGGGATACCAGTTCGTCGCGGCT 394

Qy        317 TTTCCGGACACAGTTCCGGATGGTCAGCCCGAAGCGCATCAGCAACCCGAACAATACCGG 376
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        393 TTTCCGGACACAGTTCCGGATGGTCAGCCCGAAGCGCATCAGCAACCCGAACAATACCGG 334

Qy        377 CGACAGCCGGAACTGCCGTGCCGGTGTGCAGATTAATGACAGCGGTGCGGCGCTGGGATA 436
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        333 CGACAGCCGGAACTGCCGTGCCGGTGTGCAGATTAATGACAGCGGTGCGGCGCTGGGATA 274

Qy        437 TTACGTCAGCGAGGACGGGTATCCTGGCTGGATGCCGCAGAAATGGACATGGATACCCCG 496
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        273 TTACGTCAGCGAGGACGGGTATCCTGGCTGGATGCCGCAGAAATGGACATGGATACCCCG 214

Qy        497 TGAGTTACCCGGCGGGCGCGCCTCGTTCATTCACGTTTTTGAACCCGTGGAGGACGGGCA 556
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        213 TGAGTTACCCGGCGGGCGCGCCTCGTTCATTCACGTTTTTGAACCCGTGGAGGACGGGCA 154

Qy        557 GACTCGCGGTGCAAATGTGTTTTACAGCGTGATGGAGCAGATGAAGATGCTCGACACGCT 616
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        153 GACTCGCGGTGCAAATGTGTTTTACAGCGTGATGGAGCAGATGAAGATGCTCGACACGCT 94

Qy        617 GCAGAACACGCAGCT 631
              |||||||||||||||
Db         93 GCAGAACACGCAGCT 79


RESULT 15
CB520716/c
LOCUS       CB520716                791 bp    mRNA    linear   EST 09-JUL-2003
DEFINITION  UI-M-GI0-cej-j-06-0-UI.r1 NIH_BMAP_GI0 Mus musculus cDNA clone
```

```
            IMAGE:6840391 5', mRNA sequence.
ACCESSION   CB520716
VERSION     CB520716.1  GI:29354071
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 791)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Jim Lin, University of Iowa
            cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
            cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
            DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
            Clone Distribution: Distribution information can be found at
            http://genome.uiowa.edu/distribution/mousefl.html
            This clone was contributed by the Brain Molecular Anatomy Project
            (BMAP)
            Seq primer: pYX-5.
FEATURES             Location/Qualifiers
     source          1. .791
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:6840391"
                     /tissue_type="whole brain"
                     /dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
                     /lab_host="DH10B (T1 phage resistant)"
                     /clone_lib="NIH_BMAP_GI0"
                     /note="Organ: Brain; Vector: pYX- Asc; Site_1: EcoR I;
                     Site_2: Not I; The library was constructed according
                     Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
                     1996. Denatured RNA was size fractionated on a 1% agarose
                     gel.First strand cDNA synthesis was primed with oligo-dT
                     primer containing a Not I site .Double strand cDNA was
                     size selected according to mRNA size fraction, ligated
                     with EcoR I adaptor, digested with NotI and then cloned
                     directionally into pYX-Asc vector . The library tag
                     sequence located between the Not I site and the polyA tail
                     is AGCGAGACAG. This library was created for the University
                     Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
                     Developing Mouse Nervous System', supported by National
                     Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
                     program coordinator."
ORIGIN
```

Query Match 78.4%; Score 554.4; DB 6; Length 791;
Best Local Similarity 99.6%; Pred. No. 4.8e-155;
Matches 566; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

```
Qy        45 ATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGC 104
             ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       568 ATCCAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGC 509

Qy       105 TGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCG 164
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       508 TGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCG 449

Qy       165 GCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACG 224
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       448 GCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACG 389

Qy       225 TTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTT 284
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       388 TTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTT 329

Qy       285 CAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGC 344
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       328 CAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGC 269

Qy       345 CCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTG 404
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       268 CCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTG 209

Qy       405 CAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGC 464
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       208 CAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGC 149

Qy       465 TGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTC 524
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       148 TGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTC  89

Qy       525 ATTCACGTTTTTGAACCCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTAC-AG 583
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||| || 
Db        88 ATTCACGTTTTTGAACCCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACNAG  29

Qy       584 CGTGATGGAGCAGATGAAGATGCTCGAC 611
             |||||||||||||||||||||||||||
Db        28 CGTGATGGAGCAGATGAAGATGCTCGAC   1
```

Search completed: January 26, 2005, 12:13:54
Job time : 2228.98 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: January 25, 2005, 23:20:55 ; Search time 334.971 Seconds
(without alignments)
11079.575 Million cell updates/sec

Title: US-10-826-523-40
Perfect score: 707
Sequence: 1 ggatcccatgcgtcaatttt.....aaaattgacgcatgggatcc 707

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_23Sep04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		707	100.0	707	10	ACA55353	Aca55353 Eukaryoti
2		707	100.0	4613	10	ACA55360	Aca55360 Transform
3	c	705.4	99.8	8999	10	ACA55361	Aca55361 Transform
4	c	705.4	99.8	9009	10	ACA55362	Aca55362 Transform
5	c	705.4	99.8	9012	10	ACA55363	Aca55363 Transform
6		685	96.9	4943	10	ACA55369	Aca55369 Transform
7		685	96.9	4943	10	ACA55367	Aca55367 Transform
8		681.8	96.4	4944	10	ACA55368	Aca55368 Transform
9		585.4	82.8	48502	12	ADP70043	Adp70043 Nucleic a
10		580.6	82.1	4164	4	AAS44733	Aas44733 Bacteriop
11	c	566.2	80.1	26565	6	ABS78924	Abs78924 E. coli C
12	c	566.2	80.1	26565	10	ADH80491	Adh80491 Escherich
13		566.2	80.1	46819	9	ACD19059	Acd19059 E. coli 0
14		564.6	79.9	46897	10	ADC00585	Adc00585 Enterohae
15	c	561.4	79.4	22306	9	ACD19238	Acd19238 E. coli 0
16	c	561.4	79.4	91740	10	ADC00956	Adc00956 Enterohae
17		519.2	73.4	3661	10	ACA55354	Aca55354 Transform
18	c	511	72.3	4941	10	ACA55366	Aca55366 Transform
19	c	511	72.3	4952	10	ACA55365	Aca55365 Transform
20	c	499	70.6	4951	10	ACA55364	Aca55364 Transform
21		425.8	60.2	1427	6	ABQ50926	Abq50926 Oligonucl
22	c	425.8	60.2	1427	6	ABQ50927	Abq50927 Oligonucl
23		392.6	55.5	3822	5	AAS94485	Aas94485 DNA encod
24	c	382.6	54.1	1427	6	ABQ50928	Abq50928 Oligonucl
25		382.6	54.1	1427	6	ABQ50929	Abq50929 Oligonucl
26	c	223	31.5	34063	9	ACD19083	Acd19083 E. coli 0
27	c	223	31.5	44029	10	ADC00710	Adc00710 Enterohae
28		223	31.5	49650	10	ADC00365	Adc00365 Enterohae
29		222.2	31.4	134141	6	ABN83487	Abn83487 Escherich
30	c	171.2	24.2	822	5	AAS80174	Aas80174 DNA encod
31	c	96	13.6	847	2	AAZ27249	Aaz27249 Human sec
32		78	11.0	9980	10	ACA55352	Aca55352 pIAO-P/L
33	c	73.8	10.4	5194	4	AAC85595	Aac85595 Plasmid p
34	c	73.8	10.4	5194	9	ADA09845	Ada09845 PiggyBac
35	c	73.4	10.4	5194	4	AAC85596	Aac85596 Plasmid p
36	c	73.4	10.4	5194	9	ADA09846	Ada09846 PiggyBac
37		72.8	10.3	7670	10	ACA55351	Aca55351 pIAO-P/L
38	c	67.8	9.6	2480	8	ABV76195	Abv76195 Trichoplu
39		67.8	9.6	5679	4	AAC85602	Aac85602 Helper pl
40		67.8	9.6	5679	9	ADA09857	Ada09857 Helper pl
41	c	67.8	9.6	6723	4	AAC85593	Aac85593 phsp70/op
42		67.8	9.6	6723	4	AAC85592	Aac85592 Plasmid p
43	c	67.8	9.6	6723	9	ADA09843	Ada09843 PiggyBac/
44		67.8	9.6	6723	9	ADA09842	Ada09842 Plasmid p
45	c	67.8	9.6	7560	4	AAC85599	Aac85599 Plasmid p

ALIGNMENTS

RESULT 1
ACA55353
ID ACA55353 standard; DNA; 707 BP.
XX
AC ACA55353;
XX
DT 09-JUN-2003 (first entry)
XX
DE Eukaryotic transformation vector piggyBAC ITR cartridge.
XX
KW PiggyBac; transposon; eukaryotic transformation vector; ds;
KW transformed cell; transformed embryo; transgenic; ITR cartridge.
XX
OS Synthetic.
XX
PN US2002173634-A1.
XX
PD 21-NOV-2002.
XX
PF 30-OCT-2001; 2001US-00001189.
XX
PR 31-OCT-2000; 2000US-0244677P.
PR 01-NOV-2000; 2000US-0244984P.
XX
PA (FRAS/) FRASER M J.
PA (LIXX/) LI X.
PA (BEAM/) BEAM T.
PA (HUAV/) HUA-VAN A.
XX
PI Fraser MJ, Li X, Beam T, Hua-Van A;
XX
DR WPI; 2003-352597/33.
XX
PT New DNA molecule in the transposon piggyBac, useful for transferring
PT genes into host cells or embryos for transforming the cells of embryos
PT that can be used in making transgenic organisms.
XX
PS Example 3; Fig 3(C1); 151pp; English.
XX
CC The invention describes a DNA molecule comprising at least 163
CC consecutive nucleotide base pairs of the 3' terminal region beginning at
CC the 3' terminal base pair, and at least 125 consecutive nucleotide base
CC pairs of the 5' terminal region beginning at the 5'terminal base pair, of
CC the piggyBac molecule. The region extends from the restriction site SacI

```
CC   to the end of the piggyBac molecule. The DNA molecule in the transposon
CC   piggyBac is useful for transferring genes into host cells or embryos for
CC   transforming the cells of embryos. The transformed cells or embryos are
CC   useful for developing or making transgenic organisms. This sequence
CC   represents a minimal sequence cartridge of the eukaryotic transformation
CC   vector piggyBac
XX
SQ   Sequence 707 BP; 164 A; 174 C; 208 G; 161 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 707;  DB 10;  Length 707;
  Best Local Similarity   100.0%;  Pred. No. 2.3e-213;
  Matches  707;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GGATCCCATGCGTCAATTTTACGCAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGATCCCATGCGTCAATTTTACGCAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 60

Qy         61 ATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGGCA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGGCA 120

Qy        121 TCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCGGCATGGAAAGAGTTTG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCGGCATGGAAAGAGTTTG 180

Qy        181 CCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACGTTTACCATGATGATTC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACGTTTACCATGATGATTC 240

Qy        241 GGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTTCAGGCCACCTGGGATA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTTCAGGCCACCTGGGATA 300

Qy        301 CCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGCCCGAAGCGCATCAGCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGCCCGAAGCGCATCAGCA 360

Qy        361 ACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTGCAGATTAATGACAGCG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTGCAGATTAATGACAGCG 420

Qy        421 GTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGCTGGATGCCGCAGAAAT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGCTGGATGCCGCAGAAAT 480

Qy        481 GGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTCATTCACGTTTTTGAAC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTCATTCACGTTTTTGAAC 540

Qy        541 CCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGCGTGATGGAGCAGATGA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGCGTGATGGAGCAGATGA 600

Qy        601 AGATGCTCGACACGCTGCAGAACACGCAGCTAGATTAACCCTAGAAAGATAATCATATTG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AGATGCTCGACACGCTGCAGAACACGCAGCTAGATTAACCCTAGAAAGATAATCATATTG 660

Qy        661 TGACGTACGTTAAAGATAATCATGCGTAAAATTGACGCATGGGATCC 707
              |||||||||||||||||||||||||||||||||||||||||||||
Db        661 TGACGTACGTTAAAGATAATCATGCGTAAAATTGACGCATGGGATCC 707


RESULT 2
ACA55360
ID   ACA55360 standard; DNA; 4613 BP.
XX
AC   ACA55360;
XX
DT   06-JUN-2003  (first entry)
XX
DE   Transformation vector piggyBAC related plasmid pCRII-ITR.
XX
KW   PiggyBac; transposon; eukaryotic transformation vector; ds;
KW   transformed cell; transformed embryo; transgenic; plasmid; cyclic;
KW   circular.
XX
OS   Synthetic.
XX
PN   US2002173634-A1.
XX
PD   21-NOV-2002.
XX
PF   30-OCT-2001; 2001US-00001189.
XX
PR   31-OCT-2000; 2000US-0244677P.
PR   01-NOV-2000; 2000US-0244984P.
XX
PA   (FRAS/) FRASER M J.
PA   (LIXX/) LI X.
PA   (BEAM/) BEAM T.
PA   (HUAV/) HUA-VAN A.
XX
PI   Fraser MJ,  Li X,  Beam T,  Hua-Van A;
XX
DR   WPI; 2003-352597/33.
DR   P-PSDB; ABU70357, ABU70358.
XX
PT   New DNA molecule in the transposon piggyBac, useful for transferring
PT   genes into host cells or embryos for transforming the cells of embryos
PT   that can be used in making transgenic organisms.
XX
PS   Disclosure; Fig 10(B); 151pp; English.
XX
CC   The invention describes a DNA molecule comprising at least 163
CC   consecutive nucleotide base pairs of the 3' terminal region beginning at
CC   the 3' terminal base pair, and at least 125 consecutive nucleotide base
CC   pairs of the 5' terminal region beginning at the 5'terminal base pair, of
CC   the piggyBac molecule. The region extends from the restriction site SacI
CC   to the end of the piggyBac molecule. The DNA molecule in the transposon
CC   piggyBac is useful for transferring genes into host cells or embryos for
CC   transforming the cells of embryos. The transformed cells or embryos are
CC   useful for developing or making transgenic organisms. This sequence
CC   represents plasmid used in the creation of minimal sequence eukaryotic
CC   transformation vector piggyBac
XX
SQ   Sequence 4613 BP; 1064 A; 1183 C; 1273 G; 1093 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 707;  DB 10;  Length 4613;
  Best Local Similarity   100.0%;  Pred. No. 6e-213;
  Matches  707;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GGATCCCATGCGTCAATTTTACGCAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        294 GGATCCCATGCGTCAATTTTACGCAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 353

Qy         61 ATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGGCA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        354 ATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGGCA 413

Qy        121 TCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCGGCATGGAAAGAGTTTG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        414 TCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCGGCATGGAAAGAGTTTG 473

Qy        181 CCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACGTTTACCATGATGATTC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        474 CCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACGTTTACCATGATGATTC 533

Qy        241 GGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTTCAGGCCACCTGGGATA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        534 GGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTTCAGGCCACCTGGGATA 593

Qy        301 CCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGCCCGAAGCGCATCAGCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        594 CCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGCCCGAAGCGCATCAGCA 653

Qy        361 ACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTGCAGATTAATGACAGCG 420
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        654 ACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTGCAGATTAATGACAGCG 713

Qy        421 GTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGCTGGATGCCGCAGAAAT 480
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        714 GTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGCTGGATGCCGCAGAAAT 773

Qy        481 GGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTCATTCACGTTTTTGAAC 540
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        774 GGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTCATTCACGTTTTTGAAC 833

Qy        541 CCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGCGTGATGGAGCAGATGA 600
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        834 CCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGCGTGATGGAGCAGATGA 893

Qy        601 AGATGCTCGACACGCTGCAGAACACGCAGCTAGATTAACCCTAGAAAGATAATCATATTG 660
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        894 AGATGCTCGACACGCTGCAGAACACGCAGCTAGATTAACCCTAGAAAGATAATCATATTG 953

Qy        661 TGACGTACGTTAAAGATAATCATGCGTAAAATTGACGCATGGGATCC 707
                 |||||||||||||||||||||||||||||||||||||||||||||
Db        954 TGACGTACGTTAAAGATAATCATGCGTAAAATTGACGCATGGGATCC 1000


RESULT 3
ACA55361/c
ID   ACA55361 standard; DNA; 8999 BP.
XX
AC   ACA55361;
XX
DT   06-JUN-2003  (first entry)
XX
DE   Transformation vector piggyBAC related plasmid p(PZ)-Bac-EYFP.
XX
KW   PiggyBac; transposon; eukaryotic transformation vector; ds;
KW   transformed cell; transformed embryo; transgenic; plasmid; cyclic;
KW   circular.
XX
OS   Synthetic.
XX
PN   US2002173634-A1.
XX
PD   21-NOV-2002.
XX
PF   30-OCT-2001; 2001US-00001189.
XX
PR   31-OCT-2000; 2000US-0244677P.
PR   01-NOV-2000; 2000US-0244984P.
XX
PA   (FRAS/) FRASER M J.
PA   (LIXX/) LI X.
PA   (BEAM/) BEAM T.
PA   (HUAV/) HUA-VAN A.
XX
PI   Fraser MJ, Li X, Beam T, Hua-Van A;
XX
DR   WPI; 2003-352597/33.
XX
PT   New DNA molecule in the transposon piggyBac, useful for transferring
PT   genes into host cells or embryos for transforming the cells of embryos
PT   that can be used in making transgenic organisms.
XX
PS   Example 11; Fig 12(B); 151pp; English.
XX
CC   The invention describes a DNA molecule comprising at least 163
CC   consecutive nucleotide base pairs of the 3' terminal region beginning at
CC   the 3' terminal base pair, and at least 125 consecutive nucleotide base
CC   pairs of the 5' terminal region beginning at the 5'terminal base pair, of
CC   the piggyBac molecule. The region extends from the restriction site SacI
CC   to the end of the piggyBac molecule. The DNA molecule in the transposon
CC   piggyBac is useful for transferring genes into host cells or embryos for
CC   transforming the cells of embryos. The transformed cells or embryos are
CC   useful for developing or making transgenic organisms. This sequence
CC   represents plasmid used in the creation of minimal sequence eukaryotic
CC   transformation vector piggyBac
XX
SQ   Sequence 8999 BP; 2180 A; 2416 C; 2342 G; 2061 T; 0 U; 0 Other;

  Query Match             99.8%;  Score 705.4;  DB 10;  Length 8999;
  Best Local Similarity   99.9%;  Pred. No. 2.7e-212;
  Matches  706;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GGATCCCATGCGTCAATTTTACGCAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 60
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8950 GGATCCCATGCGTCAATTTTACGCAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 8891

Qy         61 ATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGGCA 120
                 |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db       8890 ATCATATCGTCGGGTCTTTTTTCCGGGTCAGTCATCGCCCAAGCTGGCGCTATCTGGGCA 8831

Qy        121 TCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCGGCATGGAAAGAGTTTG 180
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8830 TCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCGGCATGGAAAGAGTTTG 8771

Qy        181 CCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACGTTTACCATGATGATTC 240
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8770 CCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACGTTTACCATGATGATTC 8711

Qy        241 GGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTTCAGGCCACCTGGGATA 300
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8710 GGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTTCAGGCCACCTGGGATA 8651

Qy        301 CCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGCCCGAAGCGCATCAGCA 360
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8650 CCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGCCCGAAGCGCATCAGCA 8591

Qy        361 ACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTGCAGATTAATGACAGCG 420
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8590 ACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTGCAGATTAATGACAGCG 8531

Qy        421 GTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGCTGGATGCCGCAGAAAT 480
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8530 GTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGCTGGATGCCGCAGAAAT 8471

Qy        481 GGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTCATTCACGTTTTTGAAC 540
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8470 GGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTCATTCACGTTTTTGAAC 8411

Qy        541 CCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGCGTGATGGAGCAGATGA 600
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8410 CCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGCGTGATGGAGCAGATGA 8351

Qy        601 AGATGCTCGACACGCTGCAGAACACGCAGCTAGATTAACCCTAGAAAGATAATCATATTG 660
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8350 AGATGCTCGACACGCTGCAGAACACGCAGCTAGATTAACCCTAGAAAGATAATCATATTG 8291

Qy        661 TGACGTACGTTAAAGATAATCATGCGTAAAATTGACGCATGGGATCC 707
                 |||||||||||||||||||||||||||||||||||||||||||||
Db       8290 TGACGTACGTTAAAGATAATCATGCGTAAAATTGACGCATGGGATCC 8244


RESULT 4
ACA55362/c
ID   ACA55362 standard; DNA; 9009 BP.
XX
AC   ACA55362;
XX
DT   06-JUN-2003  (first entry)
XX
DE   Transformation vector piggyBAC related plasmid p(PZ)-Bac-ECFP.
XX
KW   PiggyBac; transposon; eukaryotic transformation vector; ds;
KW   transformed cell; transformed embryo; transgenic; plasmid; cyclic;
KW   circular.
XX
OS   Synthetic.
```

XX
PN US2002173634-A1.
XX
PD 21-NOV-2002.
XX
PF 30-OCT-2001; 2001US-00001189.
XX
PR 31-OCT-2000; 2000US-0244677P.
PR 01-NOV-2000; 2000US-0244984P.
XX
PA (FRAS/) FRASER M J.
PA (LIXX/) LI X.
PA (BEAM/) BEAM T.
PA (HUAV/) HUA-VAN A.
XX
PI Fraser MJ, Li X, Beam T, Hua-Van A;
XX
DR WPI; 2003-352597/33.
XX
PT New DNA molecule in the transposon piggyBac, useful for transferring
PT genes into host cells or embryos for transforming the cells of embryos
PT that can be used in making transgenic organisms.
XX
PS Example 11; Fig 13(B); 151pp; English.
XX
CC The invention describes a DNA molecule comprising at least 163
CC consecutive nucleotide base pairs of the 3' terminal region beginning at
CC the 3' terminal base pair, and at least 125 consecutive nucleotide base
CC pairs of the 5' terminal region beginning at the 5'terminal base pair, of
CC the piggyBac molecule. The region extends from the restriction site SacI
CC to the end of the piggyBac molecule. The DNA molecule in the transposon
CC piggyBac is useful for transferring genes into host cells or embryos for
CC transforming the cells of embryos. The transformed cells or embryos are
CC useful for developing or making transgenic organisms. This sequence
CC represents plasmid used in the creation of minimal sequence eukaryotic
CC transformation vector piggyBac
XX
SQ Sequence 9009 BP; 2179 A; 2429 C; 2341 G; 2058 T; 0 U; 2 Other;

Query Match 99.8%; Score 705.4; DB 10; Length 9009;
Best Local Similarity 99.9%; Pred. No. 2.7e-212;
Matches 706; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 GGATCCCATGCGTCAATTTTACGCAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8960 GGATCCCATGCGTCAATTTTACGCAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 8901

Qy         61 ATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGGCA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8900 ATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGGCA 8841

Qy        121 TCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCGGCATGGAAAGAGTTTG 180
              ||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db       8840 TCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGGGAGGTTGAAGCGGCATGGAAAGAGTTTG 8781

Qy        181 CCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACGTTTACCATGATGATTC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8780 CCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACGTTTACCATGATGATTC 8721

Qy        241 GGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTTCAGGCCACCTGGGATA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8720 GGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTTCAGGCCACCTGGGATA 8661

Qy        301 CCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGCCCGAAGCGCATCAGCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8660 CCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGCCCGAAGCGCATCAGCA 8601

Qy        361 ACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTGCAGATTAATGACAGCG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8600 ACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTGCAGATTAATGACAGCG 8541

Qy        421 GTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGCTGGATGCCGCAGAAAT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8540 GTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGCTGGATGCCGCAGAAAT 8481

Qy        481 GGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTCATTCACGTTTTTGAAC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8480 GGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTCATTCACGTTTTTGAAC 8421

Qy        541 CCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGCGTGATGGAGCAGATGA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8420 CCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGCGTGATGGAGCAGATGA 8361

Qy        601 AGATGCTCGACACGCTGCAGAACACGCAGCTAGATTAACCCTAGAAAGATAATCATATTG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8360 AGATGCTCGACACGCTGCAGAACACGCAGCTAGATTAACCCTAGAAAGATAATCATATTG 8301

Qy        661 TGACGTACGTTAAAGATAATCATGCGTAAAATTGACGCATGGGATCC 707
              |||||||||||||||||||||||||||||||||||||||||||||
Db       8300 TGACGTACGTTAAAGATAATCATGCGTAAAATTGACGCATGGGATCC 8254
```

RESULT 5
ACA55363/c
ID ACA55363 standard; DNA; 9012 BP.
XX
AC ACA55363;
XX
DT 06-JUN-2003 (first entry)
XX
DE Transformation vector piggyBAC related plasmid p(PZ)-Bac-EGFP.
XX
KW PiggyBac; transposon; eukaryotic transformation vector; ds;
KW transformed cell; transformed embryo; transgenic; plasmid; cyclic;
KW circular.
XX
OS Synthetic.
XX
PN US2002173634-A1.
XX
PD 21-NOV-2002.
XX
PF 30-OCT-2001; 2001US-00001189.
XX
PR 31-OCT-2000; 2000US-0244677P.
PR 01-NOV-2000; 2000US-0244984P.
XX
PA (FRAS/) FRASER M J.
PA (LIXX/) LI X.
PA (BEAM/) BEAM T.
PA (HUAV/) HUA-VAN A.
XX
PI Fraser MJ, Li X, Beam T, Hua-Van A;
XX
DR WPI; 2003-352597/33.
XX
PT New DNA molecule in the transposon piggyBac, useful for transferring
PT genes into host cells or embryos for transforming the cells of embryos
PT that can be used in making transgenic organisms.
XX
PS Example 11; Fig 14(B); 151pp; English.
XX
CC The invention describes a DNA molecule comprising at least 163
CC consecutive nucleotide base pairs of the 3' terminal region beginning at
CC the 3' terminal base pair, and at least 125 consecutive nucleotide base
CC pairs of the 5' terminal region beginning at the 5'terminal base pair, of
CC the piggyBac molecule. The region extends from the restriction site SacI
CC to the end of the piggyBac molecule. The DNA molecule in the transposon
CC piggyBac is useful for transferring genes into host cells or embryos for
CC transforming the cells of embryos. The transformed cells or embryos are
CC useful for developing or making transgenic organisms. This sequence
CC represents plasmid used in the creation of minimal sequence eukaryotic
CC transformation vector piggyBac
XX
SQ Sequence 9012 BP; 2180 A; 2419 C; 2353 G; 2060 T; 0 U; 0 Other;

```
  Query Match           99.8%;  Score 705.4;  DB 10;  Length 9012;
  Best Local Similarity   99.9%;  Pred. No. 2.7e-212;
  Matches  706;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GGATCCCATGCGTCAATTTTACGCAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8963 GGATCCCATGCGTCAATTTTACGCAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 8904

Qy         61 ATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGGCA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8903 ATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGGCA 8844

Qy        121 TCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCGGCATGGAAAGAGTTTG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8843 TCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCGGCATGGAAAGAGTTTG 8784

Qy        181 CCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACGTTTACCATGATGATTC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8783 CCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACGTTTACCATGATGATTC 8724

Qy        241 GGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTTCAGGCCACCTGGGATA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8723 GGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTTCAGGCCACCTGGGATA 8664

Qy        301 CCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGCCCGAAGCGCATCAGCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8663 CCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGCCCGAAGCGCATCAGCA 8604

Qy        361 ACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTGCAGATTAATGACAGCG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8603 ACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTGCAGATTAATGACAGCG 8544

Qy        421 GTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGCTGGATGCCGCAGAAAT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8543 GTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGCTGGATGCCGCAGAAAT 8484

Qy        481 GGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTCATTCACGTTTTTGAAC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8483 GGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTCATTCACGTTTTTGAAC 8424

Qy        541 CCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGCGTGATGGAGCAGATGA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8423 CCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGCGTGATGGAGCAGATGA 8364

Qy        601 AGATGCTCGACACGCTGCAGAACACGCAGCTAGATTAACCCTAGAAAGATAATCATATTG 660
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db       8363 AGATGCTCGACACGCTGCAGAACACGCAGCTACATTAACCCTAGAAAGATAATCATATTG 8304

Qy        661 TGACGTACGTTAAAGATAATCATGCGTAAAATTGACGCATGGGATCC 707
              |||||||||||||||||||||||||||||||||||||||||||||
Db       8303 TGACGTACGTTAAAGATAATCATGCGTAAAATTGACGCATGGGATCC 8257


RESULT 6
ACA55369
ID   ACA55369 standard; DNA; 4943 BP.
XX
AC   ACA55369;
XX
DT   06-JUN-2003  (first entry)
XX
DE   Transformation vector piggyBAC related plasmid pBS-ITR-EYFP.
XX
KW   PiggyBac; transposon; eukaryotic transformation vector; ds;
KW   transformed cell; transformed embryo; transgenic; plasmid; cyclic;
KW   circular.
XX
OS   Synthetic.
XX
PN   US2002173634-A1.
XX
PD   21-NOV-2002.
XX
PF   30-OCT-2001; 2001US-00001189.
XX
PR   31-OCT-2000; 2000US-0244677P.
PR   01-NOV-2000; 2000US-0244984P.
XX
PA   (FRAS/) FRASER M J.
PA   (LIXX/) LI X.
PA   (BEAM/) BEAM T.
PA   (HUAV/) HUA-VAN A.
XX
PI   Fraser MJ,  Li X,  Beam T,  Hua-Van A;
XX
DR   WPI; 2003-352597/33.
XX
PT   New DNA molecule in the transposon piggyBac, useful for transferring
PT   genes into host cells or embryos for transforming the cells of embryos
PT   that can be used in making transgenic organisms.
XX
PS   Example 6; Fig 20(B); 151pp; English.
XX
CC   The invention describes a DNA molecule comprising at least 163
CC   consecutive nucleotide base pairs of the 3' terminal region beginning at
CC   the 3' terminal base pair, and at least 125 consecutive nucleotide base
CC   pairs of the 5' terminal region beginning at the 5'terminal base pair, of
CC   the piggyBac molecule. The region extends from the restriction site SacI
CC   to the end of the piggyBac molecule. The DNA molecule in the transposon
CC   piggyBac is useful for transferring genes into host cells or embryos for
CC   transforming the cells of embryos. The transformed cells or embryos are
CC   useful for developing or making transgenic organisms. This sequence
CC   represents plasmid used in the creation of minimal sequence eukaryotic
CC   transformation vector piggyBac
XX
SQ   Sequence 4943 BP; 1220 A; 1302 C; 1245 G; 1176 T; 0 U; 0 Other;

  Query Match           96.9%;  Score 685;  DB 10;  Length 4943;
  Best Local Similarity   99.2%;  Pred. No. 5.9e-206;
  Matches  701;  Conservative    0;  Mismatches    0;  Indels    6;  Gaps    1;

Qy          1 GGATCCCATGCGTCAATTTTACGCAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        719 GGATCCCATGCGTCAATTTTACGCAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 778

Qy         61 ATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGGCA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        779 ATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGGCA 838

Qy        121 TCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCGGCATGGAAAGAGTTTG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        839 TCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCGGCATGGAAAGAGTTTG 898

Qy        181 CCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACGTTTACCATGATGATTC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        899 CCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACGTTTACCATGATGATTC 958

Qy        241 GGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTTCAGGCCACCTGGGATA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        959 GGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTTCAGGCCACCTGGGATA 1018

Qy        301 CCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGCCCGAAGCGCATCAGCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1019 CCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGCCCGAAGCGCATCAGCA 1078

Qy        361 ACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTGCAGATTAATGACAGCG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1079 ACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTGCAGATTAATGACAGCG 1138

Qy        421 GTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGCTGGATGCCGCAGAAAT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1139 GTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGCTGGATGCCGCAGAAAT 1198

Qy        481 GGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTCATTCACGTTTTTGAAC 540
              |||||||||||||||||||||||||||||||      ||||||||||||||||||||||
Db       1199 GGACATGGATACCCCGTGAGTTACCCGGCGG------CTCGTTCATTCACGTTTTTGAAC 1252
```

```
Qy        541 CCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGCGTGATGGAGCAGATGA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1253 CCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGCGTGATGGAGCAGATGA 1312

Qy        601 AGATGCTCGACACGCTGCAGAACACGCAGCTAGATTAACCCTAGAAAGATAATCATATTG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1313 AGATGCTCGACACGCTGCAGAACACGCAGCTAGATTAACCCTAGAAAGATAATCATATTG 1372

Qy        661 TGACGTACGTTAAAGATAATCATGCGTAAAATTGACGCATGGGATCC 707
              |||||||||||||||||||||||||||||||||||||||||||||
Db       1373 TGACGTACGTTAAAGATAATCATGCGTAAAATTGACGCATGGGATCC 1419
```

RESULT 7
ACA55367
ID ACA55367 standard; DNA; 4943 BP.
XX
AC ACA55367;
XX
DT 06-JUN-2003 (first entry)
XX
DE Transformation vector piggyBAC related plasmid pBS-ITR-ECFP.
XX
KW PiggyBac; transposon; eukaryotic transformation vector; ds;
KW transformed cell; transformed embryo; transgenic; plasmid; cyclic;
KW circular.
XX
OS Synthetic.
XX
PN US2002173634-A1.
XX
PD 21-NOV-2002.
XX
PF 30-OCT-2001; 2001US-00001189.
XX
PR 31-OCT-2000; 2000US-0244677P.
PR 01-NOV-2000; 2000US-0244984P.
XX
PA (FRAS/) FRASER M J.
PA (LIXX/) LI X.
PA (BEAM/) BEAM T.
PA (HUAV/) HUA-VAN A.
XX
PI Fraser MJ, Li X, Beam T, Hua-Van A;
XX
DR WPI; 2003-352597/33.
XX
PT New DNA molecule in the transposon piggyBac, useful for transferring
PT genes into host cells or embryos for transforming the cells of embryos
PT that can be used in making transgenic organisms.
XX
PS Example 6; Fig 18(B); 151pp; English.
XX
CC The invention describes a DNA molecule comprising at least 163
CC consecutive nucleotide base pairs of the 3' terminal region beginning at
CC the 3' terminal base pair, and at least 125 consecutive nucleotide base
CC pairs of the 5' terminal region beginning at the 5'terminal base pair, of
CC the piggyBac molecule. The region extends from the restriction site SacI
CC to the end of the piggyBac molecule. The DNA molecule in the transposon
CC piggyBac is useful for transferring genes into host cells or embryos for
CC transforming the cells of embryos. The transformed cells or embryos are
CC useful for developing or making transgenic organisms. This sequence
CC represents plasmid used in the creation of minimal sequence eukaryotic
CC transformation vector piggyBac
XX
SQ Sequence 4943 BP; 1221 A; 1305 C; 1245 G; 1172 T; 0 U; 0 Other;

Query Match 96.9%; Score 685; DB 10; Length 4943;
Best Local Similarity 99.2%; Pred. No. 5.9e-206;
Matches 701; Conservative 0; Mismatches 0; Indels 6; Gaps 1;

```
Qy          1 GGATCCCATGCGTCAATTTTACGCAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        719 GGATCCCATGCGTCAATTTTACGCAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 778

Qy         61 ATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGGCA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        779 ATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGGCA 838

Qy        121 TCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCGGCATGGAAAGAGTTTG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        839 TCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCGGCATGGAAAGAGTTTG 898

Qy        181 CCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACGTTTACCATGATGATTC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        899 CCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACGTTTACCATGATGATTC 958

Qy        241 GGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTTCAGGCCACCTGGGATA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        959 GGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTTCAGGCCACCTGGGATA 1018

Qy        301 CCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGCCCGAAGCGCATCAGCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1019 CCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGCCCGAAGCGCATCAGCA 1078

Qy        361 ACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTGCAGATTAATGACAGCG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1079 ACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTGCAGATTAATGACAGCG 1138

Qy        421 GTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGCTGGATGCCGCAGAAAT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1139 GTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGCTGGATGCCGCAGAAAT 1198

Qy        481 GGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTCATTCACGTTTTTGAAC 540
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1199 GGACATGGATACCCCGTGAGTTACCCGGCGG------CTCGTTCATTCACGTTTTTGAAC 1252

Qy        541 CCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGCGTGATGGAGCAGATGA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1253 CCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGCGTGATGGAGCAGATGA 1312

Qy        601 AGATGCTCGACACGCTGCAGAACACGCAGCTAGATTAACCCTAGAAAGATAATCATATTG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1313 AGATGCTCGACACGCTGCAGAACACGCAGCTAGATTAACCCTAGAAAGATAATCATATTG 1372

Qy        661 TGACGTACGTTAAAGATAATCATGCGTAAAATTGACGCATGGGATCC 707
              |||||||||||||||||||||||||||||||||||||||||||||
Db       1373 TGACGTACGTTAAAGATAATCATGCGTAAAATTGACGCATGGGATCC 1419
```

RESULT 8
ACA55368
ID ACA55368 standard; DNA; 4944 BP.
XX
AC ACA55368;
XX
DT 06-JUN-2003 (first entry)
XX
DE Transformation vector piggyBAC related plasmid pBS-ITR-EGFP.
XX
KW PiggyBac; transposon; eukaryotic transformation vector; ds;
KW transformed cell; transformed embryo; transgenic; plasmid; cyclic;
KW circular.
XX
OS Synthetic.
XX
PN US2002173634-A1.
XX
PD 21-NOV-2002.
XX
PF 30-OCT-2001; 2001US-00001189.
XX
PR 31-OCT-2000; 2000US-0244677P.
PR 01-NOV-2000; 2000US-0244984P.
XX

```
PA   (FRAS/) FRASER M J.
PA   (LIXX/) LI X.
PA   (BEAM/) BEAM T.
PA   (HUAV/) HUA-VAN A.
XX
PI   Fraser MJ,  Li X,  Beam T,  Hua-Van A;
XX
DR   WPI; 2003-352597/33.
XX
PT   New DNA molecule in the transposon piggyBac, useful for transferring
PT   genes into host cells or embryos for transforming the cells of embryos
PT   that can be used in making transgenic organisms.
XX
PS   Example 6; Fig 19(B); 151pp; English.
XX
CC   The invention describes a DNA molecule comprising at least 163
CC   consecutive nucleotide base pairs of the 3' terminal region beginning at
CC   the 3' terminal base pair, and at least 125 consecutive nucleotide base
CC   pairs of the 5' terminal region beginning at the 5'terminal base pair, of
CC   the piggyBac molecule. The region extends from the restriction site SacI
CC   to the end of the piggyBac molecule. The DNA molecule in the transposon
CC   piggyBac is useful for transferring genes into host cells or embryos for
CC   transforming the cells of embryos. The transformed cells or embryos are
CC   useful for developing or making transgenic organisms. This sequence
CC   represents plasmid used in the creation of minimal sequence eukaryotic
CC   transformation vector piggyBac
XX
SQ   Sequence 4944 BP; 1221 A; 1302 C; 1247 G; 1174 T; 0 U; 0 Other;

  Query Match             96.4%;  Score 681.8;  DB 10;  Length 4944;
  Best Local Similarity   98.9%;  Pred. No. 6.1e-205;
  Matches  699;  Conservative    0;  Mismatches    2;  Indels    6;  Gaps    1;

Qy          1 GGATCCCATGCGTCAATTTTACGCAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        719 GGATCCCATGCGTCAATTTTACGCAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 778

Qy         61 ATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGGCA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        779 ATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGGCA 838

Qy        121 TCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCGGCATGGAAAGAGTTTG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        839 TCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCGGCATGGAAAGAGTTTG 898

Qy        181 CCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACGTTTACCATGATGATTC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        899 CCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACGTTTACCATGATGATTC 958

Qy        241 GGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTTCAGGCCACCTGGGATA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        959 GGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTTCAGGCCACCTGGGATA 1018

Qy        301 CCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGCCCGAAGCGCATCAGCA 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1019 CCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGCCCGAAGCCCATCAGCA 1078

Qy        361 ACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTGCAGATTAATGACAGCG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1079 ACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTGCAGATTAATGACAGCG 1138

Qy        421 GTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGCTGGATGCCGCAGAAAT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1139 GTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGCTGGATGCCGCAGAAAT 1198

Qy        481 GGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTCATTCACGTTTTTGAAC 540
              |||||||||||||||||||||||||||||||      |||||||||||||||||||||||
Db       1199 GGACATGGATACCCCGTGAGTTACCCGGCGG------CTCGTTCATTCACGTTTTTGAAC 1252

Qy        541 CCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGCGTGATGGAGCAGATGA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1253 CCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGCGTGATGGAGCAGATGA 1312

Qy        601 AGATGCTCGACACGCTGCAGAACACGCAGCTAGATTAACCCTAGAAAGATAATCATATTG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db       1313 AGATGCTCGACACGCTGCAGAACACGCAGCTAGATTAACCCTAGAAAGGTAATCATATTG 1372

Qy        661 TGACGTACGTTAAAGATAATCATGCGTAAAATTGACGCATGGGATCC 707
              |||||||||||||||||||||||||||||||||||||||||||||
Db       1373 TGACGTACGTTAAAGATAATCATGCGTAAAATTGACGCATGGGATCC 1419


RESULT 9
ADP70043
ID   ADP70043 standard; DNA; 48502 BP.
XX
AC   ADP70043;
XX
DT   26-AUG-2004  (first entry)
XX
DE   Nucleic acid detection-related bacteriophage lambda DNA SeqID6.
XX
KW   nucleic acid detection; nucleic acid amplification; microchip;
KW   isothermal; chip substrate; ds.
XX
OS   Bacteriophage lambda.
XX
PN   JP2004154008-A.
XX
PD   03-JUN-2004.
XX
PF   01-NOV-2002; 2002JP-00320330.
XX
PR   01-NOV-2002; 2002JP-00320330.
XX
PA   (EIKE ) EIKEN KAGAKU KK.
XX
DR   WPI; 2004-445081/42.
XX
PT   Detecting nucleic acid such as DNA, cDNA and RNA, involves performing
PT   nucleic acid amplification in microchip under isothermal conditions, and
PT   analyzing obtained amplified product on microchip.
XX
PS   Example 2; SEQ ID NO 6; 49pp; Japanese.
XX
CC   This invention relates to a novel method of detecting nucleic acid which
CC   comprises performing nucleic acid amplification in a microchip under
CC   isothermal conditions, and analysing the obtained amplified product on
CC   the microchip. The method is useful for detecting nucleic acid and
CC   enables performing a nucleic acid amplification reaction and its analysis
CC   continuously on the microchip without affecting the chip substrate. The
CC   method enables simple, reliable and cost-effective amplification,
CC   analysis and detection of nucleic acid in a microchip. The invention
CC   enables detection of trace amounts of nucleic acid with amplification
CC   efficiency. The present sequence is that of a bacteriophage lambda DNA
CC   sequence which was used in the exemplification of the invention.
XX
SQ   Sequence 48502 BP; 12334 A; 11362 C; 12820 G; 11986 T; 0 U; 0 Other;

  Query Match             82.8%;  Score 585.4;  DB 12;  Length 48502;
  Best Local Similarity   99.8%;  Pred. No. 7.1e-174;
  Matches  586;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         45 ATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGC 104
              ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3055 ATCCAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGC 3114

Qy        105 TGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCG 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3115 TGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCG 3174

Qy        165 GCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACG 224
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3175 GCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACG 3234

Qy        225 TTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTT 284
```

```
Db       3235 TTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTT 3294
Qy        285 CAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGC 344
Db       3295 CAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGC 3354
Qy        345 CCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTG 404
Db       3355 CCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTG 3414
Qy        405 CAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGC 464
Db       3415 CAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGC 3474
Qy        465 TGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTC 524
Db       3475 TGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTC 3534
Qy        525 ATTCACGTTTTTGAACCCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGC 584
Db       3535 ATTCACGTTTTTGAACCCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGC 3594
Qy        585 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCAGAACACGCAGCT 631
Db       3595 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCAGAACACGCAGCT 3641
```

RESULT 10
AAS44733
ID AAS44733 standard; DNA; 4164 BP.
XX
AC AAS44733;
XX
DT 18-DEC-2001 (first entry)
XX
DE Bacteriophage lambda full-length polynucleotide sequence #158.
XX
KW Mammal; human; rhesus monkey; baker's yeast; fission yeast; Norway rat;
KW mouse; Chinese hamster; African clawed frog; fruit fly; dog; leukaemia;
KW cancer; lymphoma; neuroblastoma; autoimmune disorder; cell proliferation;
KW nervous system disorder; inflammatory disorder; cell differentiation; ds;
KW angiogenesis; stem cell growth factor; activin; inhibin; cartilage; burn;
KW genetic disorder; bone regeneration; tendon; ligament; tissue repair;
KW cytostatic; antirheumatic; antiarthritic; vulnerary; antiinflammatory;
KW antibacterial; immunosuppressive; vasotropic; antiparkinsonian;
KW neuroprotective; osteopathic; antidiabetic; antiasthmatic; antiallergic;
KW immunostimulant; analgesic; gene therapy.
XX
OS Bacteriophage lambda.
XX
PN WO200164834-A2.
XX
PD 07-SEP-2001.
XX
PF 26-FEB-2001; 2001WO-US004926.
XX
PR 28-FEB-2000; 2000US-00515126.
PR 18-MAY-2000; 2000US-00577409.
PR 17-JUN-2000; 2000US-00597707.
PR 14-JUL-2000; 2000US-00616807.
PR 19-SEP-2000; 2000US-00664641.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
PI Xue AJ, Yang Y, Wehrman T, Wang J, Ma Y, Wang D, Chen R, Xu C;
PI Drmanac R;
XX
DR WPI; 2001-589862/66.
DR P-PSDB; AAU27833.
XX
PT Novel polypeptides and nucleic acids obtained from cDNA libraries
PT prepared from various human tissues, for diagnosis, treatment of cancer,
PT neurological, inflammatory disorders and for use in arrays for detection.
XX
PS Claim 1; SEQ ID NO 158; 153pp; English.
XX
CC Sequences AAS44576-AAS44919 represent full-length polynucleotides and
CC contig polynucleotides encoding polypeptides of the invention. The DNA
CC and protein sequences are useful for the treatment, diagnosis and
CC prevention of various types of disorder in a mammalian subject such as a
CC human, dog, monkey, mouse, hamster or rat. The disorders include cancers
CC such as leukaemia, lymphoma and neuroblastoma, autoimmune disorders such
CC as multiple sclerosis, connective tissue disease, rheumatoid arthritis,
CC diabetes mellitus, allergic rhinitis, asthma and eczema, nervous system
CC disorders such as Parkinson's disease, Alzheimer's disease, Huntington's
CC chorea, amyotrophic lateral sclerosis, spinal muscular atrophy and
CC Wernicke disease, inflammatory disorders such as nephritis, Crohn's
CC disease, ischaemia-reperfusion injury, shock, sepsis and inflammatory
CC bowel disease. The sequences exhibit activity relating to angiogenesis,
CC cell proliferation, cell differentiation, stem cell growth factor,
CC activin or inhibin. Therefore, they can be used to manipulate stem cells
CC in culture to give rise to neuroepithelial cells that can be used to
CC augment or replace cells damaged by illness, accidental damage or genetic
CC disorders. The sequences may also be used for regeneration of bone,
CC cartilage, tendons and ligaments and in tissue repair and burn healing.
CC Note: Some sequences for this patent did not form part of the printed
CC specification, but were obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 4164 BP; 906 A; 1123 C; 1307 G; 828 T; 0 U; 0 Other;

Query Match   82.1%; Score 580.6; DB 4; Length 4164;
Best Local Similarity   99.3%; Pred. No. 6.9e-173;
Matches 583; Conservative   0; Mismatches   4; Indels   0; Gaps   0;

```
Qy         45 ATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGC 104
Db        178 ATCCAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGC 237
Qy        105 TGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCG 164
Db        238 TGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCG 297
Qy        165 GCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACG 224
Db        298 GCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACG 357
Qy        225 TTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTT 284
Db        358 TTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTT 417
Qy        285 CAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGC 344
Db        418 CAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGC 477
Qy        345 CCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTG 404
Db        478 CCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTG 537
Qy        405 CAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGC 464
Db        538 CAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGC 597
Qy        465 TGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTC 524
Db        598 TATATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTC 657
Qy        525 ATTCACGTTTTTGAACCCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGC 584
Db        658 ATTCACGTTTTTGAACCCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGC 717
Qy        585 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCAGAACACGCAGCT 631
Db        718 GTGATGGAGCAGATAAAGATGCTCGACACGCTGCAGAACACGCAGCT 764
```

```
RESULT 11
ABS78924/c
ID   ABS78924 standard; DNA; 26565 BP.
XX
AC   ABS78924;
XX
DT   17-DEC-2002  (first entry)
XX
DE   E. coli CFT073 genomic sequence #91.
XX
KW   Pathogenic; Escherichia coli CFT073 infection; livestock; pyelonephritis;
KW   urinary tract infection; open reading frame; ORF; uropathogenic;
KW   antibacterial; aropathic; nephrotropic; gene; ds.
XX
OS   Escherichia coli.
XX
PN   WO200259320-A2.
XX
PD   01-AUG-2002.
XX
PF   19-OCT-2001; 2001WO-US046833.
XX
PR   19-OCT-2000; 2000US-0242412P.
XX
PA   (WISC ) WISCONSIN ALUMNI RES FOUND.
XX
PI   Blattner FR,  Welch RA,  Burland VD;
XX
DR   WPI; 2002-691532/74.
XX
PT   New DNA sequences of the pathogenic Escherichia coli CFT073 strain,
PT   useful for preventing or treating E. coli CFT073 infection in humans or
PT   livestock.
XX
PS   Claim 1; Page 276-291; 765pp; English.
XX
CC   The present invention relates to polynucleotide sequences from the genome
CC   of the pathogenic Escherichia coli strain CFT073. Almost all the
CC   sequences present in E. coli CFT073 are absent in the previously
CC   sequenced laboratory strain K-12. The polynucleotide sequences of the
CC   invention are useful for preventing, diagnosing or treating E. coli
CC   CFT073 infection in humans or livestock. The polynucleotide sequences are
CC   useful for preventing urinary tract infections and pyelonephritis.
CC   Likewise, the polypeptides encoded by the different open reading frames
CC   (ORF1-5) are useful for generating a vaccine against uropathogenic E.
CC   coli strains. ABS78834-ABS79085 represent genomic sequences from E. coli
CC   strain CFT073
XX
SQ   Sequence 26565 BP; 7042 A; 6727 C; 5778 G; 7016 T; 0 U; 2 Other;

  Query Match             80.1%;  Score 566.2;  DB 6;  Length 26565;
  Best Local Similarity   97.8%;  Pred. No. 6.5e-168;
  Matches  574;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy         45 ATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGC 104
              ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3221 ATCCAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGC 3162

Qy        105 TGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCG 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3161 TGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCG 3102

Qy        165 GCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACG 224
              |||||||||||||||||||||| ||||| ||||||||||||||||||||||||||||||
Db       3101 GCATGGAAAGAGTTTGCCGAGGACGACTGTTGCTGCATTGACGTTGAGCGAAAACGCACG 3042

Qy        225 TTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTT 284
              || |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db       3041 TTCACCATGATGATTCGGGAAGGTGTGGCCATGCATGCCTTTAACGGTGAACTGTTCGTT 2982

Qy        285 CAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGC 344
              |||||||||||||||||||  ||||||| || ||||||||||||||||||||||||||||
Db       2981 CAGGCCACCTGGGATACCAGCCCGTCGCGACTGTTCCGGACACAGTTCCGGATGGTCAGT 2922

Qy        345 CCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTG 404
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2921 CCGAAGCGTATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTG 2862

Qy        405 CAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGC 464
              |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db       2861 CAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGCTATCCTGGC 2802

Qy        465 TGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTC 524
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2801 TGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTC 2742

Qy        525 ATTCACGTTTTTGAACCCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGC 584
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db       2741 ATTCACGTTTTTGAACCCGTGGAGGACGGGCAGACCCGCGGTGCAAATGTGTTTTACAGC 2682

Qy        585 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCAGAACACGCAGCT 631
              |||||||||||||||||||||||||||||||||||||||||||||||
Db       2681 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCAGAACACGCAGCT 2635


RESULT 12
ADH80491/c
ID   ADH80491 standard; DNA; 26565 BP.
XX
AC   ADH80491;
XX
DT   22-APR-2004  (first entry)
XX
DE   Escherichia coli CFT073 genome contig #91.
XX
KW   ds; gene; Echerichia coli; CFT073; Escherichia coli CFT073 infection.
XX
OS   Escherichia coli; CFT073.
XX
PN   US2003165870-A1.
XX
PD   04-SEP-2003.
XX
PF   01-MAR-2002; 2002US-00085959.
XX
PR   01-MAR-2002; 2002US-00085959.
XX
PA   (BLAT/) BLATTNER F R.
PA   (WELC/) WELCH R A.
PA   (BURL/) BURLAND V D.
XX
PI   Blattner FR,  Welch RA,  Burland VD;
XX
DR   WPI; 2003-863698/80.
XX
PT   New nucleic acid of Escherichia coli CFT073, useful for preparing a
PT   composition for diagnosing, treating or preventing infection caused by
PT   Escherichia coli CFT073.
XX
PS   Claim 1; SEQ ID NO 91; 4pp; English.
XX
CC   The invention relates to an isolated Escherichia coli CFT073 nucleic acid
CC   molecule. The nucleic acid is useful for preparing a composition for
CC   diagnosing, treating or preventing infection caused by Escherichia coli
CC   CFT073. The present sequence represents a contig of the Escherichia coli
CC   CFT073 genome that is not present in Escherichia coli K-12.
XX
SQ   Sequence 26565 BP; 7042 A; 6727 C; 5778 G; 7016 T; 0 U; 2 Other;

  Query Match             80.1%;  Score 566.2;  DB 10;  Length 26565;
  Best Local Similarity   97.8%;  Pred. No. 6.5e-168;
  Matches  574;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy         45 ATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGC 104
```

```
Db       3221 ATCCAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGC 3162

Qy        105 TGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCG 164

Db       3161 TGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCG 3102

Qy        165 GCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACG 224

Db       3101 GCATGGAAAGAGTTTGCCGAGGACGACTGTTGCTGCATTGACGTTGAGCGAAAACGCACG 3042

Qy        225 TTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTT 284

Db       3041 TTCACCATGATGATTCGGGAAGGTGTGGCCATGCATGCCTTTAACGGTGAACTGTTCGTT 2982

Qy        285 CAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGC 344

Db       2981 CAGGCCACCTGGGATACCAGCCCGTCGCGACTGTTCCGGACACAGTTCCGGATGGTCAGT 2922

Qy        345 CCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTG 404

Db       2921 CCGAAGCGTATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTG 2862

Qy        405 CAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGC 464

Db       2861 CAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGCTATCCTGGC 2802

Qy        465 TGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTC 524

Db       2801 TGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTC 2742

Qy        525 ATTCACGTTTTTGAACCCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGC 584

Db       2741 ATTCACGTTTTTGAACCCGTGGAGGACGGGCAGACCCGCGGTGCAAATGTGTTTTACAGC 2682

Qy        585 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCAGAACACGCAGCT 631

Db       2681 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCAGAACACGCAGCT 2635
```

```
RESULT 13
ACD19059
ID   ACD19059 standard; DNA; 46819 BP.
XX
AC   ACD19059;
XX
DT   27-OCT-2003  (revised)
DT   21-AUG-2003  (first entry)
XX
DE   E. coli 0157 unique DNA sequence OZID_72.
XX
KW   OZID; ds; acute haemorrhagic colitis; haemolytic uraemic syndrome;
KW   food poisoning.
XX
OS   Escherichia coli; strain 0157:H7.
XX
PN   US2003023075-A1.
XX
PD   30-JAN-2003.
XX
PF   01-APR-2002; 2002US-00114170.
XX
PR   04-DEC-1998;   98US-0110955P.
PR   03-DEC-1999;   99US-00453702.
XX
PA   (BLAT/) BLATTNER F R.
PA   (BURL/) BURLAND V D.
PA   (PERN/) PERNA N T.
PA   (PLUN/) PLUNKETT G.
PA   (WELC/) WELCH R.
XX
PI   Blattner FR, Burland VD, Perna NT, Plunkett G, Welch R;
XX
DR   WPI; 2003-479497/45.
XX
PT   New DNA sequences from Escherichia coli strain O157:H7, useful for
PT   detecting E. coli O157:H7 in a sample, or in designing diagnostic probes
PT   which can be used to distinguish strain O157:H7 from strain K12 using
PT   molecular techniques.
XX
PS   Claim 16; SEQ ID NO 72; 33pp; English.
XX
CC   The invention relates to an isolated DNA molecule comprising an E. coli
CC   strain 0157:H7 sequence selected from a clostridial cytotoxin-like gene,
CC   a urease gene cluster, a RTX toxin-like gene cluster, a locus of
CC   enterocyte effacement and 2 genes from its associated lymphocytic phage
CC   933W (a putative serine/threonine kinase and a tail fibre gene). E. coli
CC   0157:H7 can cause food poisoning, specifically acute haemorrhagic colitis
CC   (which can develop into haemolytic uraemic syndrome). Also included are
CC   an isolated DNA molecule comprising a nucleotide sequence identical to at
CC   least 25 contiguous nucleotides contained in DNA sequences selected from
CC   ACD18988-ACD19242 (being 255 E.coli 0157 DNA sequences which are not
CC   found in E.coli K12), a recombinant DNA construction comprising the DNA
CC   above and a method for detecting E. coli 0157:H7 (ATCC 43895) in a sample
CC   (or distinguishing between 0157 and K12) using a probe derived from one
CC   of the 255 sequences. The DNA sequences are useful in detecting E. coli
CC   O157:H7 in a sample, for the early diagnosis of humans and livestock
CC   infected with O157:H7, and in designing diagnostic probes which can be
CC   used to distinguish strain O157:H7 from strain K12 using molecular
CC   techniques. The present sequence is one of the 255 E. coli 0157:H7 DNA
CC   sequence (termed OZID_1-OZID255). Note: The sequence data for this patent
CC   did not form part of the printed specification, but was obtained in
CC   electronic format directly from the USPTO at
CC   seqdata.uspto.gov/sequence.html?DocID=20030023075 (Updated on 27-OCT-2003
CC   to standardise OS field)
XX
SQ   Sequence 46819 BP; 11858 A; 10855 C; 12774 G; 11317 T; 0 U; 15 Other;
```

```
  Query Match             80.1%;  Score 566.2;  DB 9;  Length 46819;
  Best Local Similarity   97.8%;  Pred. No. 8.6e-168;
  Matches  574;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy          45 ATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGC 104

Db       14795 ATCCAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGC 14854

Qy         105 TGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCG 164

Db       14855 TGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCG 14914

Qy         165 GCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACG 224

Db       14915 GCATGGAAAGAGTTTGCCGAAGATGACTGTTGCTGCATTGACGTTGAGCGAAAACGCACG 14974

Qy         225 TTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTT 284

Db       14975 TTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTT 15034

Qy         285 CAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGC 344

Db       15035 CAGGCCACCTGGGATACCCGTCCCTCGCGACTGTTCCGGACACAGTTCCGGATGGTCAGC 15094

Qy         345 CCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTG 404

Db       15095 CCGAAGCGCATCAGCAACCCGAACAATACCAGCGACAGCCGGAACTGCCGTGCCGGTGTG 15154

Qy         405 CAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGC 464

Db       15155 CAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGC 15214

Qy         465 TGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTC 524

Db       15215 TGGATGCCGCAGAAATGGACATGGATACCCCGCGAGTTACCCGGCGGGCGCGCTTCGTTC 15274

Qy         525 ATTCACGTTTTTGAACCCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGC 584
```

```
Db      15275 ATTCACGTCTTTGAACCCGTGGAGGACGGGCAGACCCGCGGTGCAAATGTGTTTTACAGC 15334

Qy        585 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCAGAACACGCAGCT 631
              |||||||||||||||||||||||||||||||||||||||||||||||
Db      15335 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCAGAACACGCAGCT 15381


RESULT 14
ADC00585
ID   ADC00585 standard; DNA; 46897 BP.
XX
AC   ADC00585;
XX
DT   04-DEC-2003  (first entry)
XX
DE   Enterohaemorragic E. coli 0157:H7-specific nucleic acid SEQ ID NO: 630.
XX
KW   ds; gene; enterohaemorragic; anti-bacterial.
XX
OS   Escherichia coli; 0157:H7.
XX
PN   JP2002355074-A.
XX
PD   10-DEC-2002.
XX
PF   24-JAN-2002; 2002JP-00015959.
XX
PR   24-JAN-2001; 2001JP-00112010.
XX
PA   (UYTS-) UNIV TSUKUBA.
XX
DR   WPI; 2003-451640/43.
XX
PT   Enterohemorragic Escherichia coli O157:H7-specific nucleic acid molecule
PT   and a polypeptide and its use, a polypeptide, a vector and a host cell.
XX
PS   Claim 2; SEQ ID NO 630; 2067pp; Japanese.
XX
CC   The invention relates to a novel enterohaemorragic Escherichia coli
CC   O157:H7-specific nucleic acid molecule. A polynucleotide of the invention
CC   has anti-bacterial activity. The polypeptide can be used in detection
CC   and/or treatment of O157:H7 infection. The nucleotide sequence of the
CC   genome of Enterohaemorragic E coli O157:H7 was determined. The present
CC   sequence represents an E. coli 0157:H7-specific nucleic acid of the
CC   invention.
XX
SQ   Sequence 46897 BP; 11872 A; 10948 C; 12731 G; 11346 T; 0 U; 0 Other;

  Query Match             79.9%;  Score 564.6;  DB 10;  Length 46897;
  Best Local Similarity   97.6%;  Pred. No. 2.8e-167;
  Matches  573;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy          45 ATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGC 104
               ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14795 ATCCAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGC 14854

Qy         105 TGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCG 164
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14855 TGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCG 14914

Qy         165 GCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACG 224
               |||||||||||||||||| |||||||| ||||||||||||||||||||||||||||||||
Db      14915 GCATGGAAAGAGTTTGCCGAAGATGACTGTTGCTGCATTGACGTTGAGCGAAAACGCACG 14974

Qy         225 TTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTT 284
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14975 TTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTT 15034

Qy         285 CAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGC 344
               |||||||||||||||||| || | ||||| || ||||||||||||||||||||||||||||
Db      15035 CAGGCCACCTGGGATACCCGTCCCTCGCGACTGTTCCGGACACAGTTCCGGATGGTCAGC 15094

Qy         345 CCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTG 404
               |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db      15095 CCGAAGCGCATCAGCAACCCGAACAATACCAGCGACAGCCGGAACTGCCGTGCCGGTGTG 15154

Qy         405 CAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGC 464
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      15155 CAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGC 15214

Qy         465 TGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTC 524
               |||||||||||||||||||||||||||||| |||||||||||||| ||||| ||||||
Db      15215 TGGATGCCGCAGAAATGGACATGGATACCCCGCGAGTTACCCGGCGGTCGCGCTTCGTTC 15274

Qy         525 ATTCACGTTTTTGAACCCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGC 584
               |||||||| |||||||||||||||||||||||||| |||||||||||||||||||||||
Db      15275 ATTCACGTCTTTGAACCCGTGGAGGACGGGCAGACCCGCGGTGCAAATGTGTTTTACAGC 15334

Qy         585 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCAGAACACGCAGCT 631
               ||||||||||||||||||||||||||||||||||||||||||||||
Db      15335 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCAGAACACGCAGCT 15381


RESULT 15
ACD19238/c
ID   ACD19238 standard; DNA; 22306 BP.
XX
AC   ACD19238;
XX
DT   27-OCT-2003  (revised)
DT   21-AUG-2003  (first entry)
XX
DE   E. coli 0157 unique DNA sequence OZID_251.
XX
KW   OZID; ds; acute haemorrhagic colitis; haemolytic uraemic syndrome;
KW   food poisoning.
XX
OS   Escherichia coli; strain 0157:H7.
XX
PN   US2003023075-A1.
XX
PD   30-JAN-2003.
XX
PF   01-APR-2002; 2002US-00114170.
XX
PR   04-DEC-1998;   98US-0110955P.
PR   03-DEC-1999;   99US-00453702.
XX
PA   (BLAT/) BLATTNER F R.
PA   (BURL/) BURLAND V D.
PA   (PERN/) PERNA N T.
PA   (PLUN/) PLUNKETT G.
PA   (WELC/) WELCH R.
XX
PI   Blattner FR,  Burland VD,  Perna NT,  Plunkett G,  Welch R;
XX
DR   WPI; 2003-479497/45.
XX
PT   New DNA sequences from Escherichia coli strain O157:H7, useful for
PT   detecting E. coli O157:H7 in a sample, or in designing diagnostic probes
PT   which can be used to distinguish strain O157:H7 from strain K12 using
PT   molecular techniques.
XX
PS   Claim 16; SEQ ID NO 251; 33pp; English.
XX
CC   The invention relates to an isolated DNA molecule comprising an E. coli
CC   strain 0157:H7 sequence selected from a clostridial cytotoxin-like gene,
CC   a urease gene cluster, a RTX toxin-like gene cluster, a locus of
CC   enterocyte effacement and 2 genes from its associated lymphocytic phage
CC   933W (a putative serine/threonine kinase and a tail fibre gene). E. coli
CC   0157:H7 can cause food poisoning, specifically acute haemorrhagic colitis
CC   (which can develop into haemolytic uraemic syndrome). Also included are
CC   an isolated DNA molecule comprising a nucleotide sequence identical to at
CC   least 25 contiguous nucleotides contained in DNA sequences selected from
CC   ACD18988-ACD19242 (being 255 E.coli 0157 DNA sequences which are not
CC   found in E.coli K12), a recombinant DNA construction comprising the DNA
```

CC above and a method for detecting E. coli 0157:H7 (ATCC 43895) in a sample
CC (or distinguishing between 0157 and K12) using a probe derived from one
CC of the 255 sequences. The DNA sequences are useful in detecting E. coli
CC O157:H7 in a sample, for the early diagnosis of humans and livestock
CC infected with O157:H7, and in designing diagnostic probes which can be
CC used to distinguish strain O157:H7 from strain K12 using molecular
CC techniques. The present sequence is one of the 255 E. coli 0157:H7 DNA
CC sequence (termed OZID_1-OZID255). Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from the USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20030023075 (Updated on 27-OCT-2003
CC to standardise OS field)
XX
SQ Sequence 22306 BP; 4744 A; 6842 C; 5512 G; 5192 T; 0 U; 16 Other;

Query Match 79.4%; Score 561.4; DB 9; Length 22306;
Best Local Similarity 97.3%; Pred. No. 2e-166;
Matches 571; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

```
Qy        45 ATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGC 104
             ||| || ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     17193 ATCCAGTTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGC 17134

Qy       105 TGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCG 164
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     17133 TGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCG 17074

Qy       165 GCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACG 224
             |||||||||||||||||||| |||||||| ||||||||||||||||||||||||||||||
Db     17073 GCATGGAAAGAGTTTGCCGAAGATGACTGTTGCTGCATTGACGTTGAGCGAAAACGCACG 17014

Qy       225 TTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTT 284
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     17013 TTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTT 16954

Qy       285 CAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGC 344
             |||||||||||||||||| || | |||||| || ||||||||||||||||||||||||||||
Db     16953 CAGGCCACCTGGGATACCCGTCCCTCGCGACTGTTCCGGACACAGTTCCGGATGGTCAGC 16894

Qy       345 CCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTG 404
             |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db     16893 CCGAAGCGCATCAGCAACCCGAACAATACCAGCGACAGCCGGAACTGCCGTGCCGGTGTG 16834

Qy       405 CAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGC 464
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     16833 CAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGC 16774

Qy       465 TGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTC 524
             |||||||||||||||||||||||||||||| ||||||||||||||| || || ||||||
Db     16773 TGGATGCCGCAGAAATGGACATGGATACCCCGCGAGTTACCCGGCGGTCGTGCTTCGTTC 16714

Qy       525 ATTCACGTTTTTGAACCCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGC 584
             |||||||| |||||||||||||||||||||||| |||||||||||||||||||||||||
Db     16713 ATTCACGTCTTTGAACCCGTGGAGGACGGGCAGACCCGCGGTGCAAATGTGTTTTACAGC 16654

Qy       585 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCAGAACACGCAGCT 631
             |||||||||||||||||||||||||||||||||||||||||||||||
Db     16653 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCAGAACACGCAGCT 16607
```

Search completed: January 26, 2005, 03:14:34
Job time : 338.971 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: January 26, 2005, 08:26:03 ; Search time 378.34 Seconds
(without alignments)
10737.291 Million cell updates/sec

Title: US-10-826-523-40
Perfect score: 707
Sequence: 1 ggatcccatgcgtcaatttt.....aaaattgacgcatgggatcc 707

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4300275 seqs, 2872944193 residues

Total number of hits satisfying chosen parameters: 8600550

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
20: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
21: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		707	100.0	707	13	US-10-001-189-40	Sequence 40, Appl
2		707	100.0	4613	13	US-10-001-189-46	Sequence 46, Appl
3	c	707	100.0	8999	13	US-10-001-189-48	Sequence 48, Appl
4	c	707	100.0	9012	13	US-10-001-189-49	Sequence 49, Appl
5	c	707	100.0	9013	13	US-10-001-189-50	Sequence 50, Appl
6		685	96.9	4943	13	US-10-001-189-54	Sequence 54, Appl
7		685	96.9	4944	13	US-10-001-189-55	Sequence 55, Appl
8		685	96.9	4944	13	US-10-001-189-56	Sequence 56, Appl
9	c	566.2	80.1	26565	15	US-10-085-959-91	Sequence 91, Appl
10		566.2	80.1	46819	14	US-10-114-170-72	Sequence 72, Appl
11	c	561.4	79.4	22306	14	US-10-114-170-251	Sequence 251, App
12		519.2	73.4	3662	13	US-10-001-189-41	Sequence 41, Appl
13	c	511	72.3	4941	13	US-10-001-189-53	Sequence 53, Appl
14	c	511	72.3	4951	13	US-10-001-189-51	Sequence 51, Appl
15	c	511	72.3	4952	13	US-10-001-189-52	Sequence 52, Appl
16		425.8	60.2	1427	18	US-10-363-345A-37517	Sequence 37517, A
17	c	425.8	60.2	1427	18	US-10-363-345A-37518	Sequence 37518, A
18	c	382.6	54.1	1427	18	US-10-363-345A-37519	Sequence 37519, A
19		382.6	54.1	1427	18	US-10-363-345A-37520	Sequence 37520, A
20	c	223	31.5	34063	14	US-10-114-170-96	Sequence 96, Appl
21		222.2	31.4	72480	16	US-10-418-837-2	Sequence 2, Appli
22	c	96	13.6	847	10	US-09-798-889-27	Sequence 27, Appl
23	c	96	13.6	847	16	US-10-633-680-27	Sequence 27, Appl
24		78	11.0	9984	13	US-10-001-189-63	Sequence 63, Appl
25		72.8	10.3	7670	13	US-10-001-189-57	Sequence 57, Appl
26	c	40	5.7	43	13	US-10-001-189-19	Sequence 19, Appl
27		40	5.7	3089	18	US-10-425-115-49726	Sequence 49726, A
28	c	39.8	5.6	676	18	US-10-425-115-91844	Sequence 91844, A
29	c	39.8	5.6	1673	16	US-10-425-114-21174	Sequence 21174, A
30	c	39.8	5.6	1796	16	US-10-425-114-29053	Sequence 29053, A
31	c	39.8	5.6	1872	18	US-10-425-115-91842	Sequence 91842, A
32	c	39	5.5	39	13	US-10-001-189-16	Sequence 16, Appl
33		38.4	5.4	1473	16	US-10-282-122A-20134	Sequence 20134, A
34	c	36.8	5.2	4187	17	US-10-437-963-91699	Sequence 91699, A
35		36.2	5.1	1403	17	US-10-767-701-13724	Sequence 13724, A
36		36	5.1	1246	15	US-10-369-493-27114	Sequence 27114, A
37		35.4	5.0	2073	11	US-09-758-759-158	Sequence 158, App
38		35.4	5.0	96587	11	US-09-997-722-250	Sequence 250, App
39		35.4	5.0	109519	11	US-09-758-759-1	Sequence 1, Appli
40		35.2	5.0	4941	13	US-10-001-189-53	Sequence 53, Appl
41		35.2	5.0	4951	13	US-10-001-189-51	Sequence 51, Appl
42		35.2	5.0	4952	13	US-10-001-189-52	Sequence 52, Appl
43	c	35	5.0	51	9	US-09-969-852-14	Sequence 14, Appl
44	c	34.6	4.9	330	18	US-10-425-115-68281	Sequence 68281, A
45		34.4	4.9	887	16	US-10-424-599-40607	Sequence 40607, A

ALIGNMENTS

RESULT 1
US-10-001-189-40
; Sequence 40, Application US/10001189
; Publication No. US20020173634A1
; GENERAL INFORMATION:
; APPLICANT: FRASER JR., MALCOLM J.
; APPLICANT: LI, XU
; APPLICANT: BEAM, TERESA
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPOSITION USING
; TITLE OF INVENTION: MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION
; TITLE OF INVENTION: VECTOR PIGGYBAC
; FILE REFERENCE: 835910-92098
; CURRENT APPLICATION NUMBER: US/10/001,189
; CURRENT FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: 60/244,984
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/244,677
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 40
; LENGTH: 707
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: ITR Cartridge
; OTHER INFORMATION: sequence
US-10-001-189-40

Query Match 100.0%; Score 707; DB 13; Length 707;
Best Local Similarity 100.0%; Pred. No. 6.3e-236;
Matches 707; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGATCCCATGCGTCAATTTTACGCAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 60
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

```
Db          1 GGATCCCATGCGTCAATTTTACGCAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 60
Qy         61 ATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGGCA 120
Db         61 ATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGGCA 120
Qy        121 TCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCGGCATGGAAAGAGTTTG 180
Db        121 TCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCGGCATGGAAAGAGTTTG 180
Qy        181 CCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACGTTTACCATGATGATTC 240
Db        181 CCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACGTTTACCATGATGATTC 240
Qy        241 GGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTTCAGGCCACCTGGGATA 300
Db        241 GGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTTCAGGCCACCTGGGATA 300
Qy        301 CCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGCCCGAAGCGCATCAGCA 360
Db        301 CCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGCCCGAAGCGCATCAGCA 360
Qy        361 ACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTGCAGATTAATGACAGCG 420
Db        361 ACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTGCAGATTAATGACAGCG 420
Qy        421 GTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGCTGGATGCCGCAGAAAT 480
Db        421 GTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGCTGGATGCCGCAGAAAT 480
Qy        481 GGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTCATTCACGTTTTTGAAC 540
Db        481 GGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTCATTCACGTTTTTGAAC 540
Qy        541 CCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGCGTGATGGAGCAGATGA 600
Db        541 CCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGCGTGATGGAGCAGATGA 600
Qy        601 AGATGCTCGACACGCTGCAGAACACGCAGCTAGATTAACCCTAGAAAGATAATCATATTG 660
Db        601 AGATGCTCGACACGCTGCAGAACACGCAGCTAGATTAACCCTAGAAAGATAATCATATTG 660
Qy        661 TGACGTACGTTAAAGATAATCATGCGTAAAATTGACGCATGGGATCC 707
Db        661 TGACGTACGTTAAAGATAATCATGCGTAAAATTGACGCATGGGATCC 707
```

RESULT 2
US-10-001-189-46
; Sequence 46, Application US/10001189
; Publication No. US20020173634A1
; GENERAL INFORMATION:
; APPLICANT: FRASER JR., MALCOLM J.
; APPLICANT: LI, XU
; APPLICANT: BEAM, TERESA
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPOSITION USING
; TITLE OF INVENTION: MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION
; TITLE OF INVENTION: VECTOR PIGGYBAC
; FILE REFERENCE: 835910-92098
; CURRENT APPLICATION NUMBER: US/10/001,189
; CURRENT FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: 60/244,984
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/244,677
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 46
; LENGTH: 4613
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: pCRII-ITR
; OTHER INFORMATION: sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (344)..(922)
US-10-001-189-46

Query Match 100.0%; Score 707; DB 13; Length 4613;
Best Local Similarity 100.0%; Pred. No. 1.7e-235;
Matches 707; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 GGATCCCATGCGTCAATTTTACGCAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 60
Db        294 GGATCCCATGCGTCAATTTTACGCAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 353
Qy         61 ATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGGCA 120
Db        354 ATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGGCA 413
Qy        121 TCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCGGCATGGAAAGAGTTTG 180
Db        414 TCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCGGCATGGAAAGAGTTTG 473
Qy        181 CCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACGTTTACCATGATGATTC 240
Db        474 CCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACGTTTACCATGATGATTC 533
Qy        241 GGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTTCAGGCCACCTGGGATA 300
Db        534 GGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTTCAGGCCACCTGGGATA 593
Qy        301 CCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGCCCGAAGCGCATCAGCA 360
Db        594 CCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGCCCGAAGCGCATCAGCA 653
Qy        361 ACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTGCAGATTAATGACAGCG 420
Db        654 ACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTGCAGATTAATGACAGCG 713
Qy        421 GTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGCTGGATGCCGCAGAAAT 480
Db        714 GTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGCTGGATGCCGCAGAAAT 773
Qy        481 GGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTCATTCACGTTTTTGAAC 540
Db        774 GGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTCATTCACGTTTTTGAAC 833
Qy        541 CCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGCGTGATGGAGCAGATGA 600
Db        834 CCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGCGTGATGGAGCAGATGA 893
Qy        601 AGATGCTCGACACGCTGCAGAACACGCAGCTAGATTAACCCTAGAAAGATAATCATATTG 660
Db        894 AGATGCTCGACACGCTGCAGAACACGCAGCTAGATTAACCCTAGAAAGATAATCATATTG 953
Qy        661 TGACGTACGTTAAAGATAATCATGCGTAAAATTGACGCATGGGATCC 707
Db        954 TGACGTACGTTAAAGATAATCATGCGTAAAATTGACGCATGGGATCC 1000
```

RESULT 3
US-10-001-189-48/c
; Sequence 48, Application US/10001189
; Publication No. US20020173634A1
; GENERAL INFORMATION:
; APPLICANT: FRASER JR., MALCOLM J.
; APPLICANT: LI, XU
; APPLICANT: BEAM, TERESA
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPOSITION USING
; TITLE OF INVENTION: MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION
; TITLE OF INVENTION: VECTOR PIGGYBAC
; FILE REFERENCE: 835910-92098
; CURRENT APPLICATION NUMBER: US/10/001,189
; CURRENT FILING DATE: 2001-10-30

```
;  PRIOR APPLICATION NUMBER: 60/244,984
;  PRIOR FILING DATE: 2000-11-01
;  PRIOR APPLICATION NUMBER: 60/244,677
;  PRIOR FILING DATE: 2000-10-31
;  NUMBER OF SEQ ID NOS: 70
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 48
;   LENGTH: 8999
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: p(PZ)-Bac-EYFP
;   OTHER INFORMATION: sequence
US-10-001-189-48

  Query Match           100.0%;  Score 707;  DB 13;  Length 8999;
  Best Local Similarity  100.0%;  Pred. No. 2.4e-235;
  Matches  707;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GGATCCCATGCGTCAATTTTACGCAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8950 GGATCCCATGCGTCAATTTTACGCAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 8891

Qy         61 ATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGGCA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8890 ATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGGCA 8831

Qy        121 TCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCGGCATGGAAAGAGTTTG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8830 TCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCGGCATGGAAAGAGTTTG 8771

Qy        181 CCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACGTTTACCATGATGATTC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8770 CCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACGTTTACCATGATGATTC 8711

Qy        241 GGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTTCAGGCCACCTGGGATA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8710 GGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTTCAGGCCACCTGGGATA 8651

Qy        301 CCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGCCCGAAGCGCATCAGCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8650 CCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGCCCGAAGCGCATCAGCA 8591

Qy        361 ACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTGCAGATTAATGACAGCG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8590 ACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTGCAGATTAATGACAGCG 8531

Qy        421 GTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGCTGGATGCCGCAGAAAT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8530 GTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGCTGGATGCCGCAGAAAT 8471

Qy        481 GGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTCATTCACGTTTTTGAAC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8470 GGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTCATTCACGTTTTTGAAC 8411

Qy        541 CCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGCGTGATGGAGCAGATGA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8410 CCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGCGTGATGGAGCAGATGA 8351

Qy        601 AGATGCTCGACACGCTGCAGAACACGCAGCTAGATTAACCCTAGAAAGATAATCATATTG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8350 AGATGCTCGACACGCTGCAGAACACGCAGCTAGATTAACCCTAGAAAGATAATCATATTG 8291

Qy        661 TGACGTACGTTAAAGATAATCATGCGTAAAATTGACGCATGGGATCC 707
              |||||||||||||||||||||||||||||||||||||||||||||||
Db       8290 TGACGTACGTTAAAGATAATCATGCGTAAAATTGACGCATGGGATCC 8244


RESULT 4
US-10-001-189-49/c
; Sequence 49, Application US/10001189
; Publication No. US20020173634A1
; GENERAL INFORMATION:
;  APPLICANT: FRASER JR., MALCOLM J.
;  APPLICANT:  LI, XU
;  APPLICANT:  BEAM, TERESA
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPOSITION USING
;  TITLE OF INVENTION:  MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION
;  TITLE OF INVENTION:  VECTOR PIGGYBAC
;  FILE REFERENCE: 835910-92098
;  CURRENT APPLICATION NUMBER: US/10/001,189
;  CURRENT FILING DATE:  2001-10-30
;  PRIOR APPLICATION NUMBER: 60/244,984
;  PRIOR FILING DATE: 2000-11-01
;  PRIOR APPLICATION NUMBER: 60/244,677
;  PRIOR FILING DATE: 2000-10-31
;  NUMBER OF SEQ ID NOS: 70
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 49
;   LENGTH: 9012
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: p(PZ)-Bac-ECFP
;   OTHER INFORMATION: sequence
US-10-001-189-49

  Query Match           100.0%;  Score 707;  DB 13;  Length 9012;
  Best Local Similarity  100.0%;  Pred. No. 2.4e-235;
  Matches  707;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GGATCCCATGCGTCAATTTTACGCAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8963 GGATCCCATGCGTCAATTTTACGCAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 8904

Qy         61 ATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGGCA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8903 ATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGGCA 8844

Qy        121 TCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCGGCATGGAAAGAGTTTG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8843 TCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCGGCATGGAAAGAGTTTG 8784

Qy        181 CCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACGTTTACCATGATGATTC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8783 CCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACGTTTACCATGATGATTC 8724

Qy        241 GGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTTCAGGCCACCTGGGATA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8723 GGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTTCAGGCCACCTGGGATA 8664

Qy        301 CCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGCCCGAAGCGCATCAGCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8663 CCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGCCCGAAGCGCATCAGCA 8604

Qy        361 ACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTGCAGATTAATGACAGCG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8603 ACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTGCAGATTAATGACAGCG 8544

Qy        421 GTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGCTGGATGCCGCAGAAAT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8543 GTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGCTGGATGCCGCAGAAAT 8484

Qy        481 GGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTCATTCACGTTTTTGAAC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8483 GGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTCATTCACGTTTTTGAAC 8424

Qy        541 CCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGCGTGATGGAGCAGATGA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8423 CCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGCGTGATGGAGCAGATGA 8364

Qy        601 AGATGCTCGACACGCTGCAGAACACGCAGCTAGATTAACCCTAGAAAGATAATCATATTG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8363 AGATGCTCGACACGCTGCAGAACACGCAGCTAGATTAACCCTAGAAAGATAATCATATTG 8304

Qy        661 TGACGTACGTTAAAGATAATCATGCGTAAAATTGACGCATGGGATCC 707
```

```
Db        8303 TGACGTACGTTAAAGATAATCATGCGTAAAATTGACGCATGGGATCC 8257

RESULT 5
US-10-001-189-50/c
; Sequence 50, Application US/10001189
; Publication No. US20020173634A1
; GENERAL INFORMATION:
;  APPLICANT: FRASER JR., MALCOLM J.
;  APPLICANT:  LI, XU
;  APPLICANT:  BEAM, TERESA
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPOSITION USING
; TITLE OF INVENTION:  MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION
; TITLE OF INVENTION:  VECTOR PIGGYBAC
; FILE REFERENCE: 835910-92098
; CURRENT APPLICATION NUMBER: US/10/001,189
; CURRENT FILING DATE:  2001-10-30
; PRIOR APPLICATION NUMBER: 60/244,984
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/244,677
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 50
;   LENGTH: 9013
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: p(PZ)-Bac-EGFP
;   OTHER INFORMATION: sequence
US-10-001-189-50

  Query Match             100.0%;  Score 707;  DB 13;  Length 9013;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-235;
  Matches  707;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GGATCCCATGCGTCAATTTTACGCAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8964 GGATCCCATGCGTCAATTTTACGCAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 8905

Qy         61 ATCATATCGTCGGGTCTTTTTTTCCGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGGCA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8904 ATCATATCGTCGGGTCTTTTTTTCCGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGGCA 8845

Qy        121 TCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCGGCATGGAAAGAGTTTG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8844 TCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCGGCATGGAAAGAGTTTG 8785

Qy        181 CCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACGTTTACCATGATGATTC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8784 CCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACGTTTACCATGATGATTC 8725

Qy        241 GGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTTCAGGCCACCTGGGATA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8724 GGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTTCAGGCCACCTGGGATA 8665

Qy        301 CCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGCCCGAAGCGCATCAGCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8664 CCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGCCCGAAGCGCATCAGCA 8605

Qy        361 ACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTGCAGATTAATGACAGCG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8604 ACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTGCAGATTAATGACAGCG 8545

Qy        421 GTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGCTGGATGCCGCAGAAAT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8544 GTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGCTGGATGCCGCAGAAAT 8485

Qy        481 GGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTCATTCACGTTTTTGAAC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8484 GGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTCATTCACGTTTTTGAAC 8425

Qy        541 CCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGCGTGATGGAGCAGATGA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8424 CCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGCGTGATGGAGCAGATGA 8365

Qy        601 AGATGCTCGACACGCTGCAGAACACGCAGCTAGATTAACCCTAGAAAGATAATCATATTG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8364 AGATGCTCGACACGCTGCAGAACACGCAGCTAGATTAACCCTAGAAAGATAATCATATTG 8305

Qy        661 TGACGTACGTTAAAGATAATCATGCGTAAAATTGACGCATGGGATCC 707
              ||||||||||||||||||||||||||||||||||||||||||||||
Db       8304 TGACGTACGTTAAAGATAATCATGCGTAAAATTGACGCATGGGATCC 8258

RESULT 6
US-10-001-189-54
; Sequence 54, Application US/10001189
; Publication No. US20020173634A1
; GENERAL INFORMATION:
;  APPLICANT: FRASER JR., MALCOLM J.
;  APPLICANT:  LI, XU
;  APPLICANT:  BEAM, TERESA
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPOSITION USING
; TITLE OF INVENTION:  MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION
; TITLE OF INVENTION:  VECTOR PIGGYBAC
; FILE REFERENCE: 835910-92098
; CURRENT APPLICATION NUMBER: US/10/001,189
; CURRENT FILING DATE:  2001-10-30
; PRIOR APPLICATION NUMBER: 60/244,984
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/244,677
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 54
;   LENGTH: 4943
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: PBS-ITR-ECFP
;   OTHER INFORMATION: sequence
US-10-001-189-54

  Query Match             96.9%;  Score 685;  DB 13;  Length 4943;
  Best Local Similarity   99.2%;  Pred. No. 8.6e-228;
  Matches  701;  Conservative    0;  Mismatches    0;  Indels    6;  Gaps    1;

Qy          1 GGATCCCATGCGTCAATTTTACGCAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        719 GGATCCCATGCGTCAATTTTACGCAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 778

Qy         61 ATCATATCGTCGGGTCTTTTTTTCCGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGGCA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        779 ATCATATCGTCGGGTCTTTTTTTCCGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGGCA 838

Qy        121 TCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCGGCATGGAAAGAGTTTG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        839 TCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCGGCATGGAAAGAGTTTG 898

Qy        181 CCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACGTTTACCATGATGATTC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        899 CCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACGTTTACCATGATGATTC 958

Qy        241 GGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTTCAGGCCACCTGGGATA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        959 GGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTTCAGGCCACCTGGGATA 1018

Qy        301 CCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGCCCGAAGCGCATCAGCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1019 CCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGCCCGAAGCGCATCAGCA 1078

Qy        361 ACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTGCAGATTAATGACAGCG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1079 ACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTGCAGATTAATGACAGCG 1138
```

```
Qy        421 GTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGCTGGATGCCGCAGAAAT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1139 GTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGCTGGATGCCGCAGAAAT 1198

Qy        481 GGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTCATTCACGTTTTTGAAC 540
              |||||||||||||||||||||||||||||||      |||||||||||||||||||||||
Db       1199 GGACATGGATACCCCGTGAGTTACCCGGCGG------CTCGTTCATTCACGTTTTTGAAC 1252

Qy        541 CCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGCGTGATGGAGCAGATGA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1253 CCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGCGTGATGGAGCAGATGA 1312

Qy        601 AGATGCTCGACACGCTGCAGAACACGCAGCTAGATTAACCCTAGAAAGATAATCATATTG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1313 AGATGCTCGACACGCTGCAGAACACGCAGCTAGATTAACCCTAGAAAGATAATCATATTG 1372

Qy        661 TGACGTACGTTAAAGATAATCATGCGTAAAATTGACGCATGGGATCC 707
              |||||||||||||||||||||||||||||||||||||||||||||
Db       1373 TGACGTACGTTAAAGATAATCATGCGTAAAATTGACGCATGGGATCC 1419
```

RESULT 7
US-10-001-189-55
; Sequence 55, Application US/10001189
; Publication No. US20020173634A1
; GENERAL INFORMATION:
; APPLICANT: FRASER JR., MALCOLM J.
; APPLICANT: LI, XU
; APPLICANT: BEAM, TERESA
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPOSITION USING
; TITLE OF INVENTION: MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION
; TITLE OF INVENTION: VECTOR PIGGYBAC
; FILE REFERENCE: 835910-92098
; CURRENT APPLICATION NUMBER: US/10/001,189
; CURRENT FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: 60/244,984
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/244,677
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 55
; LENGTH: 4944
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: PBS-ITR-EGFP
; OTHER INFORMATION: sequence
US-10-001-189-55

Query Match 96.9%; Score 685; DB 13; Length 4944;
Best Local Similarity 99.2%; Pred. No. 8.6e-228;
Matches 701; Conservative 0; Mismatches 0; Indels 6; Gaps 1;

```
Qy          1 GGATCCCATGCGTCAATTTTACGCAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        719 GGATCCCATGCGTCAATTTTACGCAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 778

Qy         61 ATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGGCA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        779 ATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGGCA 838

Qy        121 TCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCGGCATGGAAAGAGTTTG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        839 TCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCGGCATGGAAAGAGTTTG 898

Qy        181 CCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACGTTTACCATGATGATTC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        899 CCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACGTTTACCATGATGATTC 958

Qy        241 GGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTTCAGGCCACCTGGGATA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        959 GGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTTCAGGCCACCTGGGATA 1018

Qy        301 CCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGCCCGAAGCGCATCAGCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1019 CCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGCCCGAAGCGCATCAGCA 1078

Qy        361 ACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTGCAGATTAATGACAGCG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1079 ACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTGCAGATTAATGACAGCG 1138

Qy        421 GTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGCTGGATGCCGCAGAAAT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1139 GTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGCTGGATGCCGCAGAAAT 1198

Qy        481 GGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTCATTCACGTTTTTGAAC 540
              |||||||||||||||||||||||||||||||      |||||||||||||||||||||||
Db       1199 GGACATGGATACCCCGTGAGTTACCCGGCGG------CTCGTTCATTCACGTTTTTGAAC 1252

Qy        541 CCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGCGTGATGGAGCAGATGA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1253 CCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGCGTGATGGAGCAGATGA 1312

Qy        601 AGATGCTCGACACGCTGCAGAACACGCAGCTAGATTAACCCTAGAAAGATAATCATATTG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1313 AGATGCTCGACACGCTGCAGAACACGCAGCTAGATTAACCCTAGAAAGATAATCATATTG 1372

Qy        661 TGACGTACGTTAAAGATAATCATGCGTAAAATTGACGCATGGGATCC 707
              |||||||||||||||||||||||||||||||||||||||||||||
Db       1373 TGACGTACGTTAAAGATAATCATGCGTAAAATTGACGCATGGGATCC 1419
```

RESULT 8
US-10-001-189-56
; Sequence 56, Application US/10001189
; Publication No. US20020173634A1
; GENERAL INFORMATION:
; APPLICANT: FRASER JR., MALCOLM J.
; APPLICANT: LI, XU
; APPLICANT: BEAM, TERESA
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPOSITION USING
; TITLE OF INVENTION: MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION
; TITLE OF INVENTION: VECTOR PIGGYBAC
; FILE REFERENCE: 835910-92098
; CURRENT APPLICATION NUMBER: US/10/001,189
; CURRENT FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: 60/244,984
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/244,677
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 56
; LENGTH: 4944
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: pBS-ITR-EYFP
; OTHER INFORMATION: sequence
US-10-001-189-56

Query Match 96.9%; Score 685; DB 13; Length 4944;
Best Local Similarity 99.2%; Pred. No. 8.6e-228;
Matches 701; Conservative 0; Mismatches 0; Indels 6; Gaps 1;

```
Qy          1 GGATCCCATGCGTCAATTTTACGCAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        719 GGATCCCATGCGTCAATTTTACGCAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 778

Qy         61 ATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGGCA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        779 ATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGGCA 838

Qy        121 TCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCGGCATGGAAAGAGTTTG 180
```

```
Db        839 TCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCGGCATGGAAAGAGTTTG 898
Qy        181 CCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACGTTTACCATGATGATTTC 240
Db        899 CCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACGTTTACCATGATGATTTC 958
Qy        241 GGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTTCAGGCCACCTGGGATA 300
Db        959 GGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTTCAGGCCACCTGGGATA 1018
Qy        301 CCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGCCCGAAGCGCATCAGCA 360
Db       1019 CCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGCCCGAAGCGCATCAGCA 1078
Qy        361 ACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTGCAGATTAATGACAGCG 420
Db       1079 ACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTGCAGATTAATGACAGCG 1138
Qy        421 GTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGCTGGATGCCGCAGAAAT 480
Db       1139 GTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGCTGGATGCCGCAGAAAT 1198
Qy        481 GGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTCATTCACGTTTTTGAAC 540
Db       1199 GGACATGGATACCCCGTGAGTTACCCGGCGG------CTCGTTCATTCACGTTTTTGAAC 1252
Qy        541 CCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGCGTGATGGAGCAGATGA 600
Db       1253 CCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGCGTGATGGAGCAGATGA 1312
Qy        601 AGATGCTCGACACGCTGCAGAACACGCAGCTAGATTAACCCTAGAAAGATAATCATATTG 660
Db       1313 AGATGCTCGACACGCTGCAGAACACGCAGCTAGATTAACCCTAGAAAGATAATCATATTG 1372
Qy        661 TGACGTACGTTAAAGATAATCATGCGTAAAATTGACGCATGGGATCC 707
Db       1373 TGACGTACGTTAAAGATAATCATGCGTAAAATTGACGCATGGGATCC 1419
```

RESULT 9
US-10-085-959-91/c
; Sequence 91, Application US/10085959
; Publication No. US20030165870A1
; GENERAL INFORMATION:
; APPLICANT: Blattner, Frederick R.
; APPLICANT: Welch, Rodney A.
; APPLICANT: Burland, Valerie D.
; TITLE OF INVENTION: No. US20030165870A1el Sequence of E. Coli CFT073
; FILE REFERENCE: 960296.97648
; CURRENT APPLICATION NUMBER: US/10/085,959
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: 60/242,412
; PRIOR FILING DATE: 2000-10-19
; NUMBER OF SEQ ID NOS: 255
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 91
; LENGTH: 26565
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (14853)..(14853)
; OTHER INFORMATION: Unsure
US-10-085-959-91

Query Match 80.1%; Score 566.2; DB 15; Length 26565;
Best Local Similarity 97.8%; Pred. No. 7.2e-186;
Matches 574; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

```
Qy         45 ATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGC 104
Db       3221 ATCCAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGC 3162
Qy        105 TGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCG 164
Db       3161 TGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCG 3102
Qy        165 GCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACG 224
Db       3101 GCATGGAAAGAGTTTGCCGAGGACGACTGTTGCTGCATTGACGTTGAGCGAAAACGCACG 3042
Qy        225 TTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTT 284
Db       3041 TTCACCATGATGATTCGGGAAGGTGTGGCCATGCATGCCTTTAACGGTGAACTGTTCGTT 2982
Qy        285 CAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGC 344
Db       2981 CAGGCCACCTGGGATACCAGCCCGTCGCGACTGTTCCGGACACAGTTCCGGATGGTCAGT 2922
Qy        345 CCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTG 404
Db       2921 CCGAAGCGTATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTG 2862
Qy        405 CAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGC 464
Db       2861 CAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGCTATCCTGGC 2802
Qy        465 TGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTC 524
Db       2801 TGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTC 2742
Qy        525 ATTCACGTTTTTGAACCCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGC 584
Db       2741 ATTCACGTTTTTGAACCCGTGGAGGACGGGCAGACCCGCGGTGCAAATGTGTTTTACAGC 2682
Qy        585 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCAGAACACGCAGCT 631
Db       2681 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCAGAACACGCAGCT 2635
```

RESULT 10
US-10-114-170-72
; Sequence 72, Application US/10114170
; Publication No. US20030023075A1
; GENERAL INFORMATION:
; APPLICANT: Blattner, Frederick R.
; Burland, Valerie
; Perna, Nicole T.
; Plunkett, Guy
; Welch, Rod
; TITLE OF INVENTION: No. US20030023075A1el Sequences of E. coli O157
; NUMBER OF SEQUENCES: 265
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53701-2113
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch. 1.44Mb storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 8.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/114,170
; FILING DATE: 01-Apr-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/453,702
; FILING DATE: 03-DEC-1999
; APPLICATION NUMBER: 60/110,955
; FILING DATE: 04-DEC-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J.

```
;             REGISTRATION NUMBER: 27386
;             REFERENCE/DOCKET NUMBER: 960296.95017
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (608) 251-5000
;             TELEFAX: (608) 251-9166
;   INFORMATION FOR SEQ ID NO: 72:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 46819
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;        MOLECULE TYPE: DNA  (genomic)
;        SEQUENCE DESCRIPTION: SEQ ID NO: 72:
US-10-114-170-72

  Query Match            80.1%;  Score 566.2;  DB 14;  Length 46819;
  Best Local Similarity  97.8%;  Pred. No. 9.7e-186;
  Matches  574;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy         45 ATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGC 104
              ||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14795 ATCCAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGC 14854

Qy        105 TGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCG 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14855 TGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCG 14914

Qy        165 GCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACG 224
              |||||||||||||||||||| |||||||| ||||||||||||||||||||||||||||||
Db      14915 GCATGGAAAGAGTTTGCCGAAGATGACTGTTGCTGCATTGACGTTGAGCGAAAACGCACG 14974

Qy        225 TTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTT 284
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14975 TTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTT 15034

Qy        285 CAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGC 344
              |||||||||||||||| || | ||||| || ||||||||||||||||||||||||||||||
Db      15035 CAGGCCACCTGGGATACCCGTCCCTCGCGACTGTTCCGGACACAGTTCCGGATGGTCAGC 15094

Qy        345 CCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTG 404
              |||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db      15095 CCGAAGCGCATCAGCAACCCGAACAATACCAGCGACAGCCGGAACTGCCGTGCCGGTGTG 15154

Qy        405 CAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGC 464
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      15155 CAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGC 15214

Qy        465 TGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTC 524
              |||||||||||||||||||||||||||||| |||||||||||||||||||| ||||||
Db      15215 TGGATGCCGCAGAAATGGACATGGATACCCCGCGAGTTACCCGGCGGGCGCGCTTCGTTC 15274

Qy        525 ATTCACGTTTTTGAACCCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGC 584
              |||||||| |||||||||||||||||||||||| ||||||||||||||||||||||||||
Db      15275 ATTCACGTCTTTGAACCCGTGGAGGACGGGCAGACCCGCGGTGCAAATGTGTTTTACAGC 15334

Qy        585 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCAGAACACGCAGCT 631
              |||||||||||||||||||||||||||||||||||||||||||||||
Db      15335 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCAGAACACGCAGCT 15381


RESULT 11
US-10-114-170-251/c
; Sequence 251, Application US/10114170
; Publication No. US20030023075A1
;   GENERAL INFORMATION:
;        APPLICANT: Blattner, Frederick R.
;                   Burland,     Valerie
;                   Perna,  Nicole T.
;                   Plunkett, Guy
;                   Welch,  Rod
;        TITLE OF INVENTION: No. US20030023075A1el Sequences of E. coli O157
;        NUMBER OF SEQUENCES: 265
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Quarles & Brady
;             STREET: 1 South Pinckney Street
;             CITY: Madison
;             STATE: WI
;             COUNTRY: US
;             ZIP: 53701-2113
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette, 3.50 inch. 1.44Mb storage
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: Word Perfect 8.0
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/114,170
;             FILING DATE: 01-Apr-2002
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/453,702
;             FILING DATE: 03-DEC-1999
;             APPLICATION NUMBER: 60/110,955
;             FILING DATE: 04-DEC-1998
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Seay, Nicholas J.
;             REGISTRATION NUMBER: 27386
;             REFERENCE/DOCKET NUMBER: 960296.95017
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (608) 251-5000
;             TELEFAX: (608) 251-9166
;   INFORMATION FOR SEQ ID NO: 251:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 22306
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;        MOLECULE TYPE: DNA  (genomic)
;        SEQUENCE DESCRIPTION: SEQ ID NO: 251:
US-10-114-170-251

  Query Match            79.4%;  Score 561.4;  DB 14;  Length 22306;
  Best Local Similarity  97.3%;  Pred. No. 3.1e-184;
  Matches  571;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy         45 ATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGC 104
              ||| || |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      17193 ATCCAGTTGCATCAGGATCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGC 17134

Qy        105 TGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCG 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      17133 TGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCG 17074

Qy        165 GCATGGAAAGAGTTTGCCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACG 224
              |||||||||||||||||||| |||||||| ||||||||||||||||||||||||||||||
Db      17073 GCATGGAAAGAGTTTGCCGAAGATGACTGTTGCTGCATTGACGTTGAGCGAAAACGCACG 17014

Qy        225 TTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTT 284
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      17013 TTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTT 16954

Qy        285 CAGGCCACCTGGGATACCAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGC 344
              |||||||||||||||| || | ||||| || ||||||||||||||||||||||||||||||
Db      16953 CAGGCCACCTGGGATACCCGTCCCTCGCGACTGTTCCGGACACAGTTCCGGATGGTCAGC 16894

Qy        345 CCGAAGCGCATCAGCAACCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTG 404
              |||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db      16893 CCGAAGCGCATCAGCAACCCGAACAATACCAGCGACAGCCGGAACTGCCGTGCCGGTGTG 16834

Qy        405 CAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGC 464
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      16833 CAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGC 16774

Qy        465 TGGATGCCGCAGAAATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTC 524
              |||||||||||||||||||||||||||||| |||||||||||||| || || ||||||
Db      16773 TGGATGCCGCAGAAATGGACATGGATACCCCGCGAGTTACCCGGCGGTCGTGCTTCGTTC 16714
```

```
Qy        525 ATTCACGTTTTTGAACCCGTGGAGGACGGGCAGACTCGCGGTGCAAATGTGTTTTACAGC 584
              ||||||||| |||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db      16713 ATTCACGTCTTTGAACCCGTGGAGGACGGGCAGACCCGCGGTGCAAATGTGTTTTACAGC 16654

Qy        585 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCAGAACACGCAGCT 631
              |||||||||||||||||||||||||||||||||||||||||||||||
Db      16653 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCAGAACACGCAGCT 16607


RESULT 12
US-10-001-189-41
; Sequence 41, Application US/10001189
; Publication No. US20020173634A1
; GENERAL INFORMATION:
;  APPLICANT: FRASER JR., MALCOLM J.
;  APPLICANT:  LI, XU
;  APPLICANT:  BEAM, TERESA
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPOSITION USING
;  TITLE OF INVENTION:  MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION
;  TITLE OF INVENTION:  VECTOR PIGGYBAC
;  FILE REFERENCE: 835910-92098
;  CURRENT APPLICATION NUMBER: US/10/001,189
;  CURRENT FILING DATE:  2001-10-30
;  PRIOR APPLICATION NUMBER: 60/244,984
;  PRIOR FILING DATE: 2000-11-01
;  PRIOR APPLICATION NUMBER: 60/244,677
;  PRIOR FILING DATE: 2000-10-31
;  NUMBER OF SEQ ID NOS: 70
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 41
;  LENGTH: 3662
;  TYPE: DNA
;  ORGANISM: Artificial Sequence
;  FEATURE:
;  OTHER INFORMATION: Description of Artificial Sequence: pXL-Bac
;  OTHER INFORMATION: sequence
US-10-001-189-41

  Query Match             73.4%;  Score 519.2;  DB 13;  Length 3662;
  Best Local Similarity   98.5%;  Pred. No. 6.9e-170;
  Matches  524;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          2 GATCCCATGCGTCAATTTTACGCAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGGA 61
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        983 GATCCCATGCGTCAATTTTACGCAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGGA 1042

Qy         62 TCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGGCAT 121
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1043 TCATATCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGGCAT 1102

Qy        122 CGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCGGCATGGAAAGAGTTTGC 181
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1103 CGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCGGCATGGAAAGAGTTTGC 1162

Qy        182 CGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACGTTTACCATGATGATTCG 241
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1163 CGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACGTTTACCATGATGATTCG 1222

Qy        242 GGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTTCAGGCCACCTGGGATAC 301
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1223 GGAAGGTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTTCAGGCCACCTGGGATAC 1282

Qy        302 CAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGCCCGAAGCGCATCAGCAA 361
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1283 CAGTTCGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGCCCGAAGCGCATCAGCAA 1342

Qy        362 CCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTGCAGATTAATGACAGCGG 421
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1343 CCCGAACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTGCAGATTAATGACAGCGG 1402

Qy        422 TGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGCTGGATGCCGCAGAAATG 481
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1403 TGCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGCTGGATGCCGCAGAAATG 1462

Qy        482 GACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCCTCGTTCATTCACGTT 533
              |||||||||||||||||||||||||||||||||||||||||| | |  | ||| | |
Db       1463 GACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCTTGGCGTAATCATGGT 1514


RESULT 13
US-10-001-189-53/c
; Sequence 53, Application US/10001189
; Publication No. US20020173634A1
; GENERAL INFORMATION:
;  APPLICANT: FRASER JR., MALCOLM J.
;  APPLICANT:  LI, XU
;  APPLICANT:  BEAM, TERESA
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPOSITION USING
;  TITLE OF INVENTION:  MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION
;  TITLE OF INVENTION:  VECTOR PIGGYBAC
;  FILE REFERENCE: 835910-92098
;  CURRENT APPLICATION NUMBER: US/10/001,189
;  CURRENT FILING DATE:  2001-10-30
;  PRIOR APPLICATION NUMBER: 60/244,984
;  PRIOR FILING DATE: 2000-11-01
;  PRIOR APPLICATION NUMBER: 60/244,677
;  PRIOR FILING DATE: 2000-10-31
;  NUMBER OF SEQ ID NOS: 70
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 53
;  LENGTH: 4941
;  TYPE: DNA
;  ORGANISM: Artificial Sequence
;  FEATURE:
;  OTHER INFORMATION: Description of Artificial Sequence: pXL-Bac-ECFP
;  OTHER INFORMATION: sequence
US-10-001-189-53

  Query Match             72.3%;  Score 511;  DB 13;  Length 4941;
  Best Local Similarity   100.0%;  Pred. No. 5.9e-167;
  Matches  511;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          7 CATGCGTCAATTTTACGCAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGGATCATA 66
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1129 CATGCGTCAATTTTACGCAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGGATCATA 1070

Qy         67 TCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGGCATCGGGG 126
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1069 TCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGGCATCGGGG 1010

Qy        127 AGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCGGCATGGAAAGAGTTTGCCGAGG 186
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1009 AGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCGGCATGGAAAGAGTTTGCCGAGG 950

Qy        187 ATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACGTTTACCATGATGATTCGGGAAG 246
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        949 ATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACGTTTACCATGATGATTCGGGAAG 890

Qy        247 GTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTTCAGGCCACCTGGGATACCAGTT 306
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        889 GTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTTCAGGCCACCTGGGATACCAGTT 830

Qy        307 CGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGCCCGAAGCGCATCAGCAACCCGA 366
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        829 CGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGCCCGAAGCGCATCAGCAACCCGA 770

Qy        367 ACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTGCAGATTAATGACAGCGGTGCGG 426
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        769 ACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTGCAGATTAATGACAGCGGTGCGG 710

Qy        427 CGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGCTGGATGCCGCAGAAATGGACAT 486
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        709 CGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGCTGGATGCCGCAGAAATGGACAT 650

Qy        487 GGATACCCCGTGAGTTACCCGGCGGGCGCGC 517
              |||||||||||||||||||||||||||||||
```

```
Db        649 GGATACCCCGTGAGTTACCCGGCGGGCGCGC 619


RESULT 14
US-10-001-189-51/c
; Sequence 51, Application US/10001189
; Publication No. US20020173634A1
; GENERAL INFORMATION:
;  APPLICANT: FRASER JR., MALCOLM J.
;  APPLICANT:  LI, XU
;  APPLICANT:  BEAM, TERESA
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPOSITION USING
;  TITLE OF INVENTION:  MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION
;  TITLE OF INVENTION:  VECTOR PIGGYBAC
;  FILE REFERENCE: 835910-92098
;  CURRENT APPLICATION NUMBER: US/10/001,189
;  CURRENT FILING DATE:  2001-10-30
;  PRIOR APPLICATION NUMBER: 60/244,984
;  PRIOR FILING DATE: 2000-11-01
;  PRIOR APPLICATION NUMBER: 60/244,677
;  PRIOR FILING DATE: 2000-10-31
;  NUMBER OF SEQ ID NOS: 70
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 51
;   LENGTH: 4951
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: pXL-Bac-EYFP
;   OTHER INFORMATION: sequence
US-10-001-189-51

  Query Match             72.3%;  Score 511;  DB 13;  Length 4951;
  Best Local Similarity   100.0%;  Pred. No. 5.9e-167;
  Matches  511;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          7 CATGCGTCAATTTTACGCAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGGATCATA 66
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1129 CATGCGTCAATTTTACGCAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGGATCATA 1070

Qy         67 TCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGGCATCGGGG 126
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1069 TCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGGCATCGGGG 1010

Qy        127 AGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCGGCATGGAAAGAGTTTGCCGAGG 186
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1009 AGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCGGCATGGAAAGAGTTTGCCGAGG 950

Qy        187 ATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACGTTTACCATGATGATTCGGGAAG 246
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        949 ATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACGTTTACCATGATGATTCGGGAAG 890

Qy        247 GTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTTCAGGCCACCTGGGATACCAGTT 306
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        889 GTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTTCAGGCCACCTGGGATACCAGTT 830

Qy        307 CGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGCCCGAAGCGCATCAGCAACCCGA 366
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        829 CGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGCCCGAAGCGCATCAGCAACCCGA 770

Qy        367 ACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTGCAGATTAATGACAGCGGTGCGG 426
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        769 ACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTGCAGATTAATGACAGCGGTGCGG 710

Qy        427 CGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGCTGGATGCCGCAGAAATGGACAT 486
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        709 CGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGCTGGATGCCGCAGAAATGGACAT 650

Qy        487 GGATACCCCGTGAGTTACCCGGCGGGCGCGC 517
              ||||||||||||||||||||||||||||||||
Db        649 GGATACCCCGTGAGTTACCCGGCGGGCGCGC 619


RESULT 15
US-10-001-189-52/c
; Sequence 52, Application US/10001189
; Publication No. US20020173634A1
; GENERAL INFORMATION:
;  APPLICANT: FRASER JR., MALCOLM J.
;  APPLICANT:  LI, XU
;  APPLICANT:  BEAM, TERESA
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPOSITION USING
;  TITLE OF INVENTION:  MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION
;  TITLE OF INVENTION:  VECTOR PIGGYBAC
;  FILE REFERENCE: 835910-92098
;  CURRENT APPLICATION NUMBER: US/10/001,189
;  CURRENT FILING DATE:  2001-10-30
;  PRIOR APPLICATION NUMBER: 60/244,984
;  PRIOR FILING DATE: 2000-11-01
;  PRIOR APPLICATION NUMBER: 60/244,677
;  PRIOR FILING DATE: 2000-10-31
;  NUMBER OF SEQ ID NOS: 70
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 52
;   LENGTH: 4952
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: pXL-Bac-EGFP
;   OTHER INFORMATION: sequence
US-10-001-189-52

  Query Match             72.3%;  Score 511;  DB 13;  Length 4952;
  Best Local Similarity   100.0%;  Pred. No. 5.9e-167;
  Matches  511;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          7 CATGCGTCAATTTTACGCAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGGATCATA 66
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1129 CATGCGTCAATTTTACGCAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGGATCATA 1070

Qy         67 TCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGGCATCGGGG 126
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1069 TCGTCGGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGGCATCGGGG 1010

Qy        127 AGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCGGCATGGAAAGAGTTTGCCGAGG 186
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1009 AGGAAGAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCGGCATGGAAAGAGTTTGCCGAGG 950

Qy        187 ATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACGTTTACCATGATGATTCGGGAAG 246
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        949 ATGACTGCTGCTGCATTGACGTTGAGCGAAAACGCACGTTTACCATGATGATTCGGGAAG 890

Qy        247 GTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTTCAGGCCACCTGGGATACCAGTT 306
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        889 GTGTGGCCATGCACGCCTTTAACGGTGAACTGTTCGTTCAGGCCACCTGGGATACCAGTT 830

Qy        307 CGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGCCCGAAGCGCATCAGCAACCCGA 366
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        829 CGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGCCCGAAGCGCATCAGCAACCCGA 770

Qy        367 ACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTGCAGATTAATGACAGCGGTGCGG 426
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        769 ACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTGCAGATTAATGACAGCGGTGCGG 710

Qy        427 CGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGCTGGATGCCGCAGAAATGGACAT 486
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        709 CGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGCTGGATGCCGCAGAAATGGACAT 650

Qy        487 GGATACCCCGTGAGTTACCCGGCGGGCGCGC 517
              ||||||||||||||||||||||||||||||||
Db        649 GGATACCCCGTGAGTTACCCGGCGGGCGCGC 619


Search completed: January 26, 2005, 17:51:21
Job time : 381.34 secs
```